ON DEHN FUNCTIONS OF INFINITE PRESENTATIONS OF GROUPS

R. I. GRIGORCHUK AND S. V. IVANOV

Abstract. We introduce two new types of Dehn functions of group presentations which seem more suitable (than the standard Dehn function) for infinite group presentations and prove the fundamental equivalence between the solvability of the word problem for a group presentation defined by a decidable set of defining words and the property of being computable for one of the newly introduced functions (this equivalence fails for the standard Dehn function). Elaborating on this equivalence and making use of this function, we obtain a characterization of finitely generated groups for which the word problem can be solved in nondeterministic polynomial time.

We also give upper bounds for these functions, as well as for the standard Dehn function, for two well-known periodic groups. In particular, we prove that the (standard) Dehn function of a 2-group Γ of intermediate growth, defined by a system of defining relators due to Lysenok, is bounded from above by $C_1 x^2 \log_2 x$, where $C_1 > 1$ is a constant. We also show that the (standard) Dehn function of a free m-generator Burnside group $B(m,n)$ of exponent $n \geq 2^{48}$, where n is either odd or divisible by 2^9, defined by a minimal system of defining relators, is bounded from above by the subquadratic function $x^{19/12}$.

1. Introduction

Let a finitely generated group G be defined by a presentation in terms of generators and defining relators

$$G = \langle\ \mathcal{A} \ \|\ \mathcal{R}\ \rangle\ , \tag{1}$$

where $\mathcal{A} = \{a_1, \dots, a_m\}$ is a finite alphabet and $\mathcal{R}$ is a set of defining relators which are nonempty cyclically reduced words over the alphabet $\mathcal{A}^{\pm 1} = \mathcal{A} \cup \mathcal{A}^{-1}$. Let $F(\mathcal{A})$ denote the free group over $\mathcal{A}$, $|W|$ mean the length of a word $W \in F(\mathcal{A})$ over the alphabet $\mathcal{A}^{\pm 1}$, and $\langle\langle \mathcal{R} \rangle\rangle$ denote the normal closure of $\mathcal{R}$ in $F(\mathcal{A})$. Then the notation (1) means that G is the quotient group $F(\mathcal{A})/\langle\langle \mathcal{R} \rangle\rangle$. Recall that a presentation (1) is called *finite* if $\mathcal{R}$ is finite and $\mathcal{R}$ is termed *decidable* (or recursive) if there is an algorithm to decide whether a given word over $\mathcal{A}^{\pm 1}$ belongs to $\mathcal{R}$.

Let $K(\mathcal{A}, \mathcal{R})$ be a 2-complex associated with the presentation (1) so that $K(\mathcal{A}, \mathcal{R})$ has a single 0-cell, oriented 1-cells of $K(\mathcal{A}, \mathcal{R})$ are in bijective correspondence with letters of $\mathcal{A}^{\pm 1}$, and 2-cells of $K(\mathcal{A}, \mathcal{R})$ are in bijective correspondence with the words of $\mathcal{R}$ that naturally determine the attaching maps of the 2-cells. Thus $K(\mathcal{A}, \mathcal{R})$ is the standard geometric realization of (1) with the fundamental group $\pi_1(K(\mathcal{A}, \mathcal{R}))$ being isomorphic to G.

2000 *Mathematics Subject Classification.* Primary 20E08, 20F05, 20F06, 20F10, 20F50, 20F65, 20F69.

The first named author is partially supported by NSF grants DMS 04-56185, DMS 06-00975 and by the Swiss National Science Foundation. The second named author is supported in part by NSF grant DMS 04-00746.

By a *van Kampen diagram* over the presentation (1) we mean a planar, finite, connected and simply connected 2-complex Δ which is equipped with a continuous cellular map $\mu : \Delta \mapsto K(\mathcal{A}, \mathcal{R})$ whose restriction on every cell of Δ is a homeomorphism. By an *edge* of Δ we mean the closure of a 1-cell. If e is an oriented edge of Δ and $\mu(e)$ corresponds to a letter $a \in \mathcal{A}^{\pm 1}$, then a is termed the *label* of e and is denoted $\varphi(e)$. Note $\varphi(e^{-1}) = \varphi(e)^{-1}$, where e^{-1} denotes the edge with opposite orientation. If $p = e_1 \dots e_k$ is a path in Δ, where $e_1, \dots, e_k$ are oriented edges of Δ, then we set $\varphi(p) = \varphi(e_1) \dots \varphi(e_k)$. According to a well-known lemma of van Kampen, see [23, 30], a word W belongs to $\langle\langle \mathcal{R} \rangle\rangle$ if and only if there exists a van Kampen diagram over (1) whose boundary path $\partial\Delta$ is labeled by the word W, in which case we write $\varphi(\partial\Delta) \equiv W$, where the sign $\equiv$ means the literal equality of (cyclic) words. The number of j-cells of a van Kampen diagram Δ is denoted by $|\Delta(j)|$, $j \in \{0, 1, 2\}$.

Let $W \in \langle\langle \mathcal{R} \rangle\rangle$ and $j \in \{0, 1, 2\}$. Define $L_j(W)$ to be the minimal number of j-cells in a van Kampen diagram Δ over (1) whose boundary $\partial\Delta$ is labeled by the cyclic word W, that is

$$L_j(W) = \min\{\, |\Delta(j)| \mid \varphi(\partial\Delta) \equiv W \} \ . \tag{2}$$

For an integer $x \geq 1$, define

$$f_j(x) = \max\{\, L_j(W) \mid W \in \langle\langle \mathcal{R} \rangle\rangle \text{ and } |W| \leq x \} \ . \tag{3}$$

Recall that if $W \in \langle\langle \mathcal{R} \rangle\rangle$, then, in $F(\mathcal{A})$, one has an equality of the form

$$W = \prod_{i=1}^{L} X_i R_i^{\varepsilon_i} X_i^{-1} \ , \tag{4}$$

where $X_i \in F(\mathcal{A})$, $R_i \in \mathcal{R}$, $\varepsilon_i \in \{\pm 1\}$ and where we allow $L = 0$ (when $W = 1$ in $F(\mathcal{A})$). An equivalent way to define the number $L_2(W)$ is to pick minimal $L \geq 0$ over all products of the form (4).

When $j = 2$, the foregoing definition (3) defines the well-known *Dehn function* $f_2(x)$ of a group presentation (1) that has been subject of intensive research for the past twenty years. Recall that the concept of the Dehn function $f_2(x)$ of a group presentation (1) was introduced into group theory by Gromov in his seminal article [14] in 1987 and now is fundamental in geometric group theory. For instance, a finite presentation (1) defines a (word) hyperbolic group if and only if its Dehn function $f_2(x)$ is bounded from above by a linear function. For a finite presentation (1), the solvability of the word problem is equivalent to the property of being computable for the function $f_2(x)$. As was pointed out by the referee, in 1985, Madlener and Otto [26] considered a notion of the "derivational complexity" which was an earlier version of the definition of the Dehn function and which was defined by means of regarding a presentation of a monoid, or a group, as a string rewriting system.

The object of this paper is to introduce and study the functions $f_0(x)$, $f_1(x)$, to establish basic relations between the solvability of the word problem for presentation (1) and the property of being computable for functions $f_0(x)$, $f_1(x)$, and to give upper bounds for the functions $f_j(x)$ in case of presentations of two well-known periodic (or torsion) groups. The function $f_j(x)$, $j \in \{0, 1, 2\}$, will be referred to as the *Dehn j-function* of a group presentation (1) (with the prefix "j-" frequently omitted).

As was pointed out above, for a finite presentation (1), the solvability of the word problem is equivalent to the property of being computable for the Dehn function

$f_2(x)$. However, if $\mathcal{R}$ is not finite, then this equivalence breaks down and generates a few interesting problems which we state for two Dehn functions $f_2(x)$ and $f_0(x)$.

Problem 1.1. *Let the relator set $\mathcal{R}$ of a presentation* (1) *be decidable and $j \in \{0, 2\}$. Prove or disprove that*

(a) *If the word problem for* (1) *is solvable, then the Dehn j-function $f_j(x)$ of* (1) *is computable.*

(b) *If the Dehn j-function $f_j(x)$ of* (1) *is computable, then the word problem for* (1) *is solvable.*

Remarkably, the Dehn 1-function $f_1(x)$ can be used in place of $f_2(x)$ to fix the failing equivalence, see Example 2.4. Taking advantage of counting 1-cells in place of 2-cells, we have the following basic result.

Theorem 1.2. *Let $\mathcal{R}$ in* (1) *be decidable. Then the word problem for* (1) *is solvable if and only if the Dehn* 1*-function $f_1(x)$ of* (1) *is computable.*

Elaborating on this equivalence, we obtain a characterization of finitely generated groups for which the word problem could be solved in nondeterministic polynomial time.

Theorem 1.3. *Let a group $G = \langle a_1, \dots, a_m \rangle$ be generated by elements $a_1, \dots, a_m$. Then the word problem for G is in* **NP**, *i.e. it can be solved algorithmically in nondeterministic polynomial time, if and only if there exists a presentation $\langle a_1, \dots, a_m \,\|\, \mathcal{R} \rangle$ for G such that its Dehn* 1*-function $f_1(x)$ is bounded by a polynomial and the problem to decide whether a word $W = W(a_1^{\pm 1}, \dots, a_m^{\pm 1})$ belongs to $\mathcal{R}$ is in* **NP**.

Let us emphasize that the complexity, as well as the solvability, of the word problem for a finitely generated group G, given by a presentation (1), does not actually depend on $\mathcal{R}$. Indeed, we need to decide whether or not a given word W over $\mathcal{A}^{\pm 1}$ represents the identity element of G, i.e. whether $W \in \langle\langle \mathcal{R} \rangle\rangle$. On the other hand, when speaking about the word problem for a finitely generated group G, we always assume that a finite generating set, say $\mathcal{A}$, is fixed for G, i.e. G is regarded as the quotient group $F(\mathcal{A})/\langle\langle \mathcal{R} \rangle\rangle$ for some $\mathcal{R}$, see, for example, Theorem 1.3 above and Corollary 1.11 below.

Theorem 1.3 is reminiscent of a deep result of Birget, Ol'shanskii, Rips and Sapir [5] that states that a finitely generated group G has the word problem in **NP** if and only if G is isomorphic to a subgroup of a finitely presented group whose Dehn function is bounded by a polynomial. However, unlike the result of [5], Theorem 1.3 is a straightforward corollary of the advantageous definition of the function $f_1(x)$. Unlike the result of [5], Theorem 1.3 also holds for other computational classes, for example, it holds with **PSPACE** in place of **NP** (recall that the class **PSPACE** consists of decision problems that could be solved in polynomial space, for more details see [36]).

Let $\mathcal{F}$ denote the set of functions $f : \mathbb{N} \to \mathbb{N} \cup \{0\}$, where $\mathbb{N} = \{1, 2, \dots\}$ is the set of natural numbers. If $f, g \in \mathcal{F}$, we write $f \preccurlyeq g$ if there is an integer $C > 0$ such that $f(x) \le Cg(Cx) + Cx$ for every $x \in \mathbb{N}$. Two functions f, $g \in \mathcal{F}$ are said to be (linearly) equivalent, denoted $f \simeq g$, if $f \preccurlyeq g$ and $g \preccurlyeq f$. It is evident that this relation $\simeq$ is indeed an equivalence relation and the equivalence class $[f]_\simeq$ could be regarded as the growth rate of a function $f \in \mathcal{F}$.

The natural questions on relations between functions $f_j(x)$, $j = 0, 1, 2$, are addressed in the following.

Theorem 1.4. (a) *For every presentation* (1), $f_0(x) \le 2f_1(x)$ *and* $f_2(x) \le 2f_1(x)$. *In particular,* $f_0(x) \preccurlyeq f_1(x)$ *and* $f_2(x) \preccurlyeq f_1(x)$.

(b) *Let every element in* $\mathcal{R}$ *of* (1) *be a word of length* > 1. *Then the Dehn functions* $f_0(x)$ *and* $f_1(x)$ *of* (1) *are equivalent.*

Making use of the ideas of proofs of Theorems 1.2 and 1.4(b), we will give a positive solution to Problem 1.1 in a special case for $j = 0$.

Theorem 1.5. *Let* $\mathcal{R}$ *in* (1) *be decidable and, for every* $R \in \mathcal{R}$, $|R| > 1$. *Then the word problem for* (1) *is solvable if and only if the Dehn* 0*-function* $f_0(x)$ *of* (1) *is computable. Moreover,* $f_0(x)$ *is computable if and only if* $f_1(x)$ *is computable.*

To generalize the Andrews–Curtis and Magnus conjectures, the following operation over a finite presentation (1) is introduced in the article [22]: An element of $\mathcal{R}$ is replaced by another element if doing so does not change the normal closure of $\mathcal{R}$. More generally, consider replacement of a finite subset $\mathcal{S}$ of $\mathcal{R}$ by another finite set $\mathcal{U}$ of cyclically reduced words if doing so does not change the normal closure of $\mathcal{R}$. As in [22], call this operation a *T-transformation.* Another operation over a presentation (1), called *stabilization*, is to add/delete a letter b to/from both $\mathcal{A}$ and $\mathcal{R}$ (when deleting b from $\mathcal{A}$ and $\mathcal{R}$, $b^{\pm 1}$ must not occur in any other word of $\mathcal{R}$). Note that both T-transformation and stabilization are special cases of Tietze transformations, see [27, Section 1.5].

Theorem 1.6. (a) *Let* $\langle \mathcal{A} \parallel \mathcal{R}' \rangle$ *be obtained from a presentation* (1) *by a T-transformation and* $f'_j(x)$, $f_j(x)$, $j = 0, 1, 2$, *be their corresponding Dehn functions. Then* $\quad f'_0(x) \preccurlyeq f_1(x), \;\; f'_1(x) \simeq f_1(x), \;\; f'_2(x) \simeq f_2(x)$.

(b) *Let* $\langle \mathcal{A}' \parallel \mathcal{R}' \rangle$ *be obtained from a presentation* (1) *by a stabilization and* $f'_j(x)$, $f_j(x)$, $j = 0, 1, 2$, *be their corresponding Dehn functions. Then* $f'_j(x) \simeq f_j(x)$ *for* $j = 0, 1, 2$.

Making use of T-transformations and stabilizations, we will prove the equivalence between all of the Dehn j-functions $f_j(x)$, $j = 0, 1, 2$, for a finite presentation.

Theorem 1.7. *Let* (1) *be a finite presentation. Then all of its Dehn functions* $f_0(x)$, $f_1(x)$, $f_2(x)$ *are equivalent.*

In Section 2, we will give Examples 2.1–2.2 of group presentations for which the pairs $(f_1(x), f_2(x))$, $(f_0(x), f_1(x))$, $(f_0(x), f_2(x))$ consist of nonequivalent functions. Hence, in general, $f_1(x) \not\simeq f_2(x)$, $f_0(x) \not\simeq f_1(x)$, $f_0(x) \not\simeq f_2(x)$. In Section 2, we will also consider Examples 2.3–2.4 of group presentations that address parts (a)–(b) of Problem 1.1 in the case $j = 2$.

It is well known (e.g. see [2], [26]) that if $\mathcal{P}$, $\mathcal{P}'$ are two finite presentations of a group G and $f_2(x)$, $f'_2(x)$ are their corresponding Dehn 2-functions, then $f_2(x) \simeq f'_2(x)$ (this also follows from Theorem 1.6(a)–(b) because $\mathcal{P}'$ can be obtained from $\mathcal{P}$ by stabilizations and T-transformations, see [27, Section 1.5]). Furthermore, if H is a subgroup of finite index of a group G that has a finite presentation (or, more generally, G and H are abstractly commensurable), then the Dehn 2-functions of finite presentations of G and H are equivalent. We also recall that, if the Dehn 2-function of a finite presentation of a group G is subquadratic, then the

group G is word hyperbolic, see [14, 31]. For more results on Dehn 2-functions see [4, 6, 5, 15, 16, 32, 33, 35] and the references cited there.

On the other hand, if we allow infinite relator sets $\mathcal{R}$ in (1), then the growth rates $[f_j(x)]_{\simeq}$, $j = 0, 1, 2$, of the Dehn j-function of G are no longer independent of a presentation of G. Indeed, if $\mathcal{R}$ consists of all cyclically reduced words of $\langle\langle \mathcal{R} \rangle\rangle$, then $f_2(x) \le 1$, $f_0(x) \le x$, $f_1(x) \le x$, and hence $f_j(x) \simeq x$, $j = 0, 1, 2$.

In general, an inclusion $\mathcal{R}' \subseteq \mathcal{R}$, provided $\langle\langle \mathcal{R}' \rangle\rangle = \langle\langle \mathcal{R} \rangle\rangle$, easily implies the inequalities $f_j(x) \le f'_j(x)$, $j = 0, 1, 2$, for the corresponding Dehn j-functions (cf. Theorem 1.6(a)) and, for this reason, it is more natural to consider presentations (1) with *minimal* $\mathcal{R}$, i.e., if $\mathcal{R}' \subseteq \mathcal{R}$ and $\langle\langle \mathcal{R}' \rangle\rangle = \langle\langle \mathcal{R} \rangle\rangle$ then $\mathcal{R}' = \mathcal{R}$. However, Y. de Conrulier (private communication) pointed out to us that there are finitely generated groups that possess no presentations (1) with minimal $\mathcal{R}$.

Investigation of Dehn j-functions $f_j(x)$, $j = 0, 1, 2$, of a presentation (1), where the relator set $\mathcal{R}$ need not be finite, seems to be an interesting and important problem, especially in the case when $\mathcal{R}$ is minimal. In this article, we obtain two results in this direction for infinite presentations of periodic groups investigated by the authors in earlier articles.

Let Γ denote the 2-group of intermediate growth that was originally discovered by the first author in [7] and later investigated in [8], [24] and other papers, see also [17, Chapter VIII]. It was shown by Lysenok [24] that the group Γ can be defined by the following presentation in terms of generators and defining relators

$$\begin{aligned} \Gamma = \langle\ a, b, c, d \parallel a^2,\ b^2,\ c^2,\ d^2,\ bcd,\ \sigma^i((ad)^4), \\ \sigma^i((adacac)^4),\ i \ge 0\ \rangle\ =\ \langle\ a, b, c, d \parallel \mathcal{R}(\infty)\ \rangle\ , \end{aligned} \tag{5}$$

where it is assumed that $\sigma^0 = \mathrm{id}$ is the identity map and that σ is the endomorphism of the free group $F(a, b, c, d)$ with the basis $\{a, b, c, d\}$ defined by

$$\sigma = \begin{cases} a \mapsto aca\ , \\ b \mapsto d\ , \\ c \mapsto b\ , \\ d \mapsto c\ . \end{cases} \tag{6}$$

The group Γ has no finite presentation, see [8, 10], and, as is shown in [10], the set $\mathcal{R}(\infty)$ of Lysenok's relators is minimal.

Theorem 1.8. *The Dehn j-functions $f_{j,\Gamma}(x)$, $j = 0, 1, 2$, of the Lysenok presentation* (5) *of the* 2*-group Γ are bounded from above by $C_1 x^2 \log_2 x$, where $C_1 > 1$ is a constant.*

It is of interest to point out that, when proving Theorem 1.8, we will construct a larger relator set $\mathcal{R}^*(\infty)$, containing $\mathcal{R}(\infty)$ (so $\mathcal{R}^*(\infty)$ is not minimal), and obtain different upper bounds for the functions $f^*_{j,\Gamma}(x)$, $j \in \{1, 2\}$, corresponding to $\mathcal{R}^*(\infty)$: $f^*_{1,\Gamma}(x) < C^*_1 x^2 \log_2 x$ and $f^*_{2,\Gamma}(x) < C^*_1 x^2$, where $C^*_1 > 1$ is a constant.

As in article [9], consider a group defined by the following presentation

$$\Gamma_t = \langle\ a, b, c, d, t \parallel \mathcal{R}(\infty),\ \ tgt^{-1}\sigma(g)^{-1},\ g \in \{a, b, c, d\}\ \rangle\ .$$

Since $\sigma(R) \in \langle\langle \mathcal{R}(\infty) \rangle\rangle$ for every $R \in \mathcal{R}(\infty)$, it follows that Γ_t is an ascending HNN-extension of Γ with the stable letter t. Thanks to the form of Lysenok's

relators $\mathcal{R}(\infty)$, the group Γ_t can also be defined by the following finite presentation

$$\Gamma_t = \langle\ a,b,c,d,t \parallel a^2,\ b^2,\ c^2,\ d^2,\ bcd,\ (ad)^4,\ (adacac)^4, \\ tgt^{-1}\sigma(g)^{-1},\ g \in \{a,b,c,d\}\ \rangle\ . \tag{7}$$

It is immediate from the definition that Γ_t is a finitely presented torsion-by-cyclic group, see also [21], [34] for more examples of finitely presented torsion-by-cyclic groups. Furthermore, as was observed in [9], the group Γ_t is amenable but not elementary amenable, a property shared by Γ. As a consequence of Theorem 1.8, we will obtain

Corollary 1.9. *The Dehn j-functions $f_{j,\Gamma_t}(x)$, $j = 0,1,2$, of the finite presentation* (7) *are bounded from above by $C_2 2^x x$, where $C_2 > 1$ is a constant.*

Recall that a free m-generator Burnside group $B(m,n)$ of exponent n is the quotient F/F^n, where F is a free group of rank m and $F^n = \langle W^n \mid W \in F\rangle$.

To construct a presentation for $B(m,n)$, as in [19, 29], we consider a total order $\preceq$ on the set of all words over the alphabet $\mathcal{A}^{\pm 1} = \{a_1^{\pm 1}, \dots, a_m^{\pm 1}\}$ such that $U \preceq V$ when $|U| \le |V|$. As above, let $F(\mathcal{A})$ denote the free group with the basis $\mathcal{A}$. Set $B(m,n,0) = F(\mathcal{A}) = \langle \mathcal{A} \parallel \emptyset\rangle$. Proceeding by induction on $i \ge 1$, assume that the group presentation $B(m,n,i-1)$ is already constructed. Let A_i be the minimal (if it exits), relative to the order $\preceq$, word over $\mathcal{A}^{\pm 1}$ such that the image of A_i has infinite order in the group defined by $B(m,n,i-1)$. Note that A_i may not exist and then our inductive process terminates and results in the presentation $B(m,n,i-1)$. If A_i does exist, then the presentation $B(m,n,i)$ is obtained from $B(m,n,i-1)$ by addition of the relator A_i^n. Clearly,

$$B(m,n,i) = \langle\ \mathcal{A} \parallel A_1^n,\ A_2^n,\ \dots,\ A_i^n\ \rangle \tag{8}$$

and, if A_i exists for every $i \ge 1$, then, taking the limit, we obtain

$$B(m,n,\infty) = \langle\ \mathcal{A} \parallel A_1^n,\ A_2^n,\ \dots,\ A_i^n,\ \dots\ \rangle\ . \tag{9}$$

Assume that $n \ge 2^{48}$ and either n is odd or divisible by 2^9. Under this assumption, it is proved in [19, Theorem B], see also [18], [20], that A_i does exist[1] for every $i \ge 1$ and the limit group $B(m,n,\infty)$ is naturally isomorphic to the free Burnside group $B(m,n) = F(\mathcal{A})/F(\mathcal{A})^n$ with the basis $\mathcal{A}$. Furthermore, it is shown in [19], see Theorem B and Lemma 21.1, that the relator set $\{A_1^n,\ A_2^n, \dots, A_i^n, \dots\}$ is decidable and minimal. We remark that, for odd $n > 10^{10}$, these results are due to Ol'shanskii [29, 30], compare with Novikov–Adian's [28], Adian's [1], and Lysenok's [25] presentations and results on $B(m,n)$.

For free Burnside groups, we prove

Theorem 1.10. *Let $m \ge 2$, $n \ge 2^{48}$ and n be either odd or divisible by 2^9. Then the Dehn* 1*-function $f_{1,B}(x)$ of the presentation* (9) *of a free m-generator Burnside group $B(m,n)$ of exponent n is bounded from above by the subquadratic function $x^{19/12}$. In addition, $f_{0,B}(x) \le 2x^{19/12}$ and $f_{2,B}(x) \le \frac{2}{n}x^{19/12}$. Furthermore, for every $i \ge 0$, the same upper bounds hold for the Dehn j-functions $f_{j,B(i)}(x)$ of the finite presentation $B(m,n,i)$ defined by* (8).

[1] In the case when $B(m,n)$ is finite, the word A_i will fail to exist for some i. We do not know whether it is possible that $B(m,n)$ is infinite and A_i does not exist for some i.

Note that the cubic upper bound $f_{2,B}(x) \le 6(n^{-1}x)^3$ for the Dehn function $f_{2,B}(x)$ of $B(m,n,\infty)$, in the case of odd $n > 10^{10}$, is due to Storozhev [30, Section 28.2]. We also mention that a linear bound $f_{2,B(i)}(x) \le K_i x$ for the Dehn function $f_{2,B(i)}(x)$ of the finite presentation $B(m,n,i)$ was obtained in [19, Lemma 21.1]. This linear upper bound implies that, for every $i \ge 0$, the group $B(m,n,i)$ is word hyperbolic. However, the constant K_i, as a function of i, grows exponentially and so the bounds $f_{2,B(i)}(x) \le K_i x$, $i = 1, 2, \dots$, do not shed any light upon an upper bound for $f_{2,B}(x)$. On the other hand, in view of results of Gromov [14] and Ol'shanskii [31] on the hyperbolicity of finitely presented groups with subquadratic Dehn 2-function, the subquadratic bound $f_{2,B(i)}(x) \le \frac{2}{n} x^{19/12}$ of Theorem 1.10 implies that, for every $i \ge 0$, the group given by presentation (8) is word hyperbolic and hence, by Theorem 1.7, its Dehn j-functions $f_{j,B(i)}$ are bounded by a linear function.

As a consequence of Theorem 1.10 and lemmas of [19], we will also derive

Corollary 1.11. *Let $n \ge 2^{48}$ and either n be odd or divisible by 2^9. Then the word and conjugacy problems for the free Burnside group $B(m,n) = F(\mathcal{A})/F(\mathcal{A})^n$ are in* **NP**.

For odd $n > 10^{10}$, Corollary 1.11 could be derived from the Storozhev's cubic bound for $f_{2,B}(x)$ and lemmas of [29]. It would be interesting to further investigate the complexity of the word problem for $B(m,n)$ and find out whether it is in **P**, i.e. it is solvable in deterministic polynomial time, or in **coNP** or, perhaps, **NP**-complete, see [36]. We also remark that the word problem for the group Γ, see (5), is known to be in **P**. In fact, the word problem for Γ can be solved in subquadratic time ($\sim y \log_2 y$)), see [10]. A deterministic algorithm for the conjugacy problem for Γ, whose running time is at least exponential, as well as the history of the question and further references could be found in [12]. It would also be desirable to make further progress on the following.

Problem 1.12. *Obtain nontrivial lower bounds for the Dehn j-functions $f_{j,\Gamma}$, $f_{j,B}$, $j \in \{1,2\}$, of presentations* (5), (9), *improve on upper bounds for $f_{j,\Gamma}$, $f_{j,B}$, and determine the growth rates $[f_{j,\Gamma}]_{\simeq}$, $[f_{j,B}]_{\simeq}$.*

Note that, by Theorem 1.4(b), $f_{0,\Gamma} \simeq f_{1,\Gamma}$ and $f_{0,B} \simeq f_{1,B}$.

An interesting notion of the verbal Dehn function $f_w(x)$ of a variety of groups, defined by a single identity $w = 1$, was introduced and investigated by Ol'shanskii and Sapir [32]. To give the definition, we let $F_\infty = F(a_1, a_2, \dots)$ be the free group over the countably infinite alphabet $\mathcal{A}_\infty = \{a_1, a_2, \dots\}$, w be a word, and $w(F_\infty)$ be the w-verbal subgroup of F_∞ generated by all values of w on F_∞. For every $U \in w(F_\infty)$, consider a product in F_∞ of the form

$$U = \prod_{j=1}^{L} X_j w(Y_{1j}, \dots, Y_{kj})^{\varepsilon_j} X_j^{-1} \ ,$$

where $X_j, Y_{ij} \in F_\infty$ and $\varepsilon_j = \pm 1$. Taking the minimal sum $\sum_{j=1}^{L}(|Y_{1j}| + \dots + |Y_{kj}|)$ over all such products for U, we obtain a number $s(U)$. Then the *verbal Dehn function* for the word w is defined by $f_w(x) = \max\{s(U) \mid |U| \le x,\ U \in w(F_\infty)\}$.

According to [32], in an unpublished work, for odd $n > 10^{10}$, Mikhailov gave an upper bound $x^{1+\varepsilon_n}$, where $\varepsilon_n > 0$ and $\varepsilon_n \to 0$ as $n \to \infty$, for the verbal Dehn function $f_{z^n}(x)$ of the Burnside variety of groups of exponent n. We conjecture that

an analogous upper bound holds for the function $f_{1,B}(x)$, where $n \gg 1$ is defined as in Theorem 1.10. However, it remains unclear whether there are any relations between the functions $f_{z^n}(x)$ and $f_{1,B}(x)$ and whether an upper bound for $f_{z^n}(x)$ could possibly yield any bound for $f_{1,B}(x)$.

As is pointed out in [32], the verbal Dehn function $f_w(x)$ is *superadditive* for every word w, i.e. $f_w(x)$ satisfies the inequality $f_w(x_1 + x_2) \ge f_w(x_1) + f_w(x_2)$. Recall that the superadditive closure $\bar f(x)$ of a function $f(x)$, $x \in \mathbb{N}$, is $\bar f(x) = \max\{f(x_1)+\cdots+f(x_r)\}$ over all sums $x = x_1+\cdots+x_r$. It is immediate that $\bar f(x)$ is superadditive. A conjecture, put forward by Guba and Sapir [16], claims that the Dehn function $f_2(x)$ of a finite presentation is equivalent to its superadditive closure $\bar f_2(x)$. It seems to be of interest to ask the same question for Dehn j-functions $f_j(x)$, $j = 0, 1, 2$, defined for an arbitrary presentation (1), that is, to ask whether $f_j(x) \simeq \bar f_j(x)$, $j = 0, 1, 2$. Speaking of the equivalence $f_j(x) \simeq \bar f_j(x)$, we remark that we do not know whether functions $f_{j,B}(x)$, $f_{j,\Gamma}(x)$ are equivalent to their superadditive closures. We do not know either whether two functions $f_{1,B}(x)$, corresponding to two different choices of defining words A_1^n, $A_2^n, \dots$ in (9), are equivalent, to leave alone the equivalence of functions $f_{1,B}(x)$ of free Burnside groups $B(m, n)$ of different ranks $m \ge 2$ and exponents $n \gg 1$.

2. Proofs of Theorems 1.2–1.7

Proof of Theorem 1.2. Suppose that the word problem is solvable for a presentation (1) with a decidable set $\mathcal{R}$. Then the normal closure $\langle\langle \mathcal{R} \rangle\rangle$ of $\mathcal{R}$ is a decidable subset of $F(\mathcal{A})$. To compute the value of $f_1(x)$ for a given integer $x \ge 1$, consider the set $\mathcal{W}_x$ of all words over the alphabet $\mathcal{A}^{\pm 1}$ of length $\le x$. For each $U \in \mathcal{W}_x$, we determine whether $U \in \langle\langle \mathcal{R} \rangle\rangle$. If so, we construct a van Kampen diagram Δ_U over (1) with $\varphi(\partial \Delta_U) \equiv U$. Now we check all van Kampen diagrams Δ such that $\varphi(\partial\Delta) \equiv U$ and $|\Delta(1)| \le |\Delta_U(1)|$ and find a diagram $\Delta_{U,1}$ over (1) with the minimal number $|\Delta_{U,1}(1)|$. Then $f_1(x) = \max\{|\Delta_{U,1}(1)| = L_1(U) \mid U \in \mathcal{W}_x \cap \langle\langle \mathcal{R} \rangle\rangle\}$ and therefore $f_1(x)$ is computable.

Conversely, suppose that the function $f_1(x)$ of (1) is computable. Let W be a word of length $|W| \le x$. Then $W \in \langle\langle \mathcal{R} \rangle\rangle$ if and only if there exists a van Kampen diagram Δ_W over (1) such that $\varphi(\partial\Delta_W) \equiv W$ and $|\Delta_W(1)| \le f_1(x)$. Note that if Π is a face (= closure of a 2-cell) in Δ_W then $|\partial\Pi| \le 2|\Delta_W(1)|$, where $|\partial\Pi|$ is the perimeter of Π. Since $\mathcal{R}$ is decidable, we can write down all words $R \in \mathcal{R}$ with $|R| \le 2f_1(x)$ and hence we can construct all possible diagrams Δ over (1) with $|\Delta(1)| \le f_1(x)$ to determine whether $W \in \langle\langle \mathcal{R} \rangle\rangle$. This proves that the set $\langle\langle \mathcal{R} \rangle\rangle$ is decidable and the word problem is solvable for (1). $\square$

Proof of Theorem 1.3. Suppose that $G = \langle a_1, \dots, a_m \rangle$ is a group generated by elements $a_1, \dots, a_m$ and the problem to decide whether a given word over the alphabet $\mathcal{A}^{\pm 1} = \{a_1^{\pm 1}, \dots, a_m^{\pm 1}\}$, represents the identity element of G is in **NP**. Consider the presentation $G = \langle \mathcal{A} \,\|\, \mathcal{R} \rangle$, where $\mathcal{R}$ consists of all nonempty cyclically reduced words R over $\mathcal{A}^{\pm 1}$ that represent the identity element G. Then, by the definitions, the problem to determine whether $U \in \mathcal{R}$ is in **NP** and the corresponding function $f_1(x)$ is bounded by the polynomial x.

Conversely, suppose that G is defined by a presentation $G = \langle \mathcal{A} \,\|\, \mathcal{R} \rangle$ such that the corresponding function $f_1(x)$ is bounded by a polynomial $p(x)$ and the problem to decide whether a word U belongs to $\mathcal{R}$ is in **NP**. For definiteness, assume that this

problem can be solved in nondeterministic time bounded by a polynomial $q(|U|)$. We need to show that the word problem for $G = \langle \mathcal{A} \,\|\, \mathcal{R} \rangle$ is in **NP**.

A word U over $\mathcal{A}^{\pm 1}$ represents the identity element G if and only if there exists a van Kampen diagram Δ_U such that $\varphi(\partial \Delta_U) \equiv U$ and $|\Delta_U(1)| \le f_1(|U|) \le p(|U|)$. We also remark that $|\Delta_U(2)| \le 2|\Delta_U(1)| \le 2p(|U|)$ and, for every face Π of Δ_U, we have $|\partial \Pi| \le 2|\Delta_U(1)| \le 2p(|U|)$. Therefore, a van Kampen diagram Δ, where $|\Delta(1)| \le p(x)$, $|\Delta(2)| \le 2p(x)$ and $|\partial \Pi| \le 2p(x)$ for every face Π in Δ, can be used as a certificate to verify that a given word U with $|U| \le x$ represents the identity element of G. This can be done in time bounded by the polynomial

$$|\Delta(2)| \cdot q(|\partial \Pi|) + |\partial \Delta| \le 2p(x) \cdot q(2p(x)) + x \ ,$$

where $q(|\partial \Pi|)$ estimates the time needed to verify that the word $\varphi(\partial \Pi)$ is in $\mathcal{R}$ and $|\partial \Delta|$ bounds the time needed to check that $\varphi(\partial \Delta) \equiv U$.

We remark that the above argument is retained with **PSPACE** in place of **NP** (recall that **NPSPACE** = **PSPACE**, see [36]). □

Proof of Theorem 1.4. Part (a). Let Δ be a van Kampen diagram over (1). Assuming that $|\Delta(1)| > 0$, it is easy to see that $|\Delta(0)| \le 2|\Delta(1)|$ and $|\Delta(2)| \le 2|\Delta(1)|$. In the notation (2), these inequalities mean that $L_0(W) \le 2L_1(W)$ and $L_2(W) \le 2L_1(W)$. Maximizing over all $W \in \langle\langle \mathcal{R} \rangle\rangle$ with $|W| \le x$, we obtain $f_0(x) \le 2f_1(x)$ and $f_2(x) \le 2f_1(x)$, as required.

Part (b). Let Δ be a van Kampen diagram over (1), $\varphi(\partial \Delta) \equiv U$ and $|\Delta(1)|$ is minimal over all diagrams Δ_1 such that $\varphi(\partial \Delta_1) \equiv U$ and $|\Delta_1(0)| = |\Delta(0)|$. To simplify the notation, let $V = |\Delta(0)|$, $E = |\Delta(1)|$, $F = |\Delta(2)|$. Let F_2 and F_3 denote the numbers of faces in Δ that have 2 and ≥ 3, respectively, edges in their boundaries. Recall that $\mathcal{R}$ has no words of length 1, whence $F = F_2 + F_3$. We also consider a planar 2-complex Δ' obtained from Δ by identifying e, f for every pair of (oriented) edges e, f such that ef^{-1} is the boundary cycle of a face of Δ. Let V', E', F', F_2', F_3' be defined for Δ' in the same manner as the numbers V, E, F, F_2, F_3 were defined for Δ. Clearly, $F_2' = 0$ and

$$E = E' + F_2 \ . \tag{10}$$

Suppose that e_1, e_2 are edges in Δ such that $(e_1)_- = (e_2)_-$ and $(e_1)_+ = (e_2)_+$ (perhaps, $(e_1)_- = (e_2)_+$), where e_- denotes the initial vertex of an oriented edge e and e_+ is the terminal vertex of e. Also, assume that $\varphi(e_1) = \varphi(e_2)$ and the subdiagram, bounded by the closed path $e_1(e_2)^{-1}$, has no other vertices than $(e_1)_-$, $(e_2)_+$. If $e_1 \ne e_2$, then one could make a surgery on Δ that would identify e_1, e_2, decrease $|\Delta(1)|$ and preserve both $|\Delta(0)|$ and $\varphi(\partial \Delta)$. Since Δ is minimal relative to $|\Delta(1)|$, it follows that the inequality $e_1 \ne e_2$ is impossible and so $e_1 = e_2$ for such a pair e_1, e_2. This remark implies that

$$F_2 \le 2mE' \ , \tag{11}$$

where m is the number of letters in $\mathcal{A}$. By the Euler formula applied to Δ', we have $V' - E' + F' = 1$ or $V - E' + F_3 = 1$, because $V' = V$, $F_3' = F_3$ and $F_2' = 0$. Note that $F_3' \le \frac{2E'}{3}$, hence $V - E' + \frac{2E'}{3} \ge 1$ and $\frac{E'}{3} \le V$. By (10)–(11), we obtain

$$|\Delta(1)| = E = E' + F_2 \le (1+2m)E' \le 3(1+2m)V = 3(1+2m)|\Delta(0)| \ . \tag{12}$$

This estimate, together with the minimality of Δ, implies that

$$f_1(x) \le 3(1+2m)f_0(x) \ .$$

Hence, $f_1(x) \preccurlyeq f_0(x)$ which, together with proven part (a), proves the equivalence of functions $f_0(x)$ and $f_1(x)$. $\square$

Proof of Theorem 1.5. Recall that if e is an oriented edge of a diagram Δ, then e_- denotes the initial vertex of e and e_+ is the terminal vertex of e. We will say that a van Kampen diagram Δ over (1) is 1-*regular* if Δ has the following property. If e_1, e_2 are edges in Δ such that $(e_1)_- = (e_2)_-$ and $(e_1)_+ = (e_2)_+$ (perhaps, $(e_1)_- = (e_2)_+$), $\varphi(e_1) = \varphi(e_2)$ and the subdiagram Δ_e with $\partial\Delta_e = e_1 e_2^{-1}$ has no other vertices than $(e_1)_-$, $(e_2)_+$, then $e_1 = e_2$. Observe that if Δ is not 1-regular and e_1, e_2 are the edges in Δ that violate the above property, then one could make a surgery on Δ, called 1-*reduction*, that would take Δ_e out of Δ and identify e_1, e_2. Note that a 1-reduction preserves both $\varphi(\partial\Delta)$ and $|\Delta(0)|$ and it decreases $|\Delta(1)|$. Hence, application of finitely many 1-reductions to Δ will yield a 1-regular diagram Δ_1 such that $\varphi(\partial\Delta_1) \equiv \varphi(\partial\Delta)$ and $|\Delta_1(0)| = |\Delta(0)|$. In particular, without loss of generality, we may assume that if $U \in \langle\langle \mathcal{R} \rangle\rangle$ and Δ is a van Kampen diagram over (1) such that $\varphi(\partial\Delta) \equiv U$ and $|\Delta(0)| = L_0(U)$, see (2), then Δ is 1-regular.

From now on assume that $|R| > 1$ for every $R \in \mathcal{R}$. Repeating the arguments of the proof of Theorem 1.4(b) aimed to prove the inequality (12), we can analogously show that if Δ is a 1-regular van Kampen diagram over (1), then

$$|\Delta(1)| \le 3(1+2m)|\Delta(0)| \ . \tag{13}$$

Now suppose that the word problem is solvable for (1) and $\mathcal{R}$ is decidable. Then $\langle\langle \mathcal{R} \rangle\rangle$ is also a decidable subset of $F(\mathcal{A})$ and, as in the proof of Theorem 1.2, for every $U \in \langle\langle \mathcal{R} \rangle\rangle$ we can effectively construct a van Kampen diagram Δ_U over (1) with $\varphi(\partial\Delta_U) \equiv U$. Applying 1-reductions to Δ_U if necessary, we may assume that Δ_U is 1-regular. By inequality (13), $|\Delta_U(1)| \le 3(1+2m)|\Delta_U(0)|$. Now let $\Delta_{U,0}$ be a van Kampen diagram over (1) such that $\varphi(\partial\Delta_{U,0}) \equiv U$ and $|\Delta_{U,0}(0)| = L_0(U)$. As was pointed out above, we may assume that $\Delta_{U,0}$ is 1-regular, hence, by (13), $|\Delta_{U,0}(1)| \le 3(1+2m)|\Delta_{U,0}(0)| \le 3(1+2m)|\Delta_U(0)|$. This means that by checking all diagrams Δ that satisfy $|\Delta(1)| \le 3(1+2m)|\Delta_U(0)|$ we can compute the number $L_0(U)$. Thus $f_0(x) = \max\{L_0(U) \mid U \in \langle\langle \mathcal{R} \rangle\rangle,\ |U| \le x\}$ is also computable.

Conversely, suppose that the function $f_0(x)$ is computable for (1) and let W be a word with $|W| \le x$. Then $W \in \langle\langle \mathcal{R} \rangle\rangle$ if and only if there exists a 1-regular van Kampen diagram Δ_W over (1) such that $\varphi(\partial\Delta_W) \equiv W$ and $|\Delta_W(0)| \le f_0(x)$. In view of inequality (13),

$$|\Delta_W(1)| \le 3(1+2m)|\Delta_W(0)| \le 3(1+2m) f_0(x) \ . \tag{14}$$

Hence, as in the proof of Theorem 1.2, we can construct all possible diagrams Δ over (1) with $|\Delta(1)| \le 3(1+2m)f_0(x)$ and determine whether or not $W \in \langle\langle \mathcal{R} \rangle\rangle$. This proves that the word problem is solvable for (1).

Finally, by Theorem 1.2, the Dehn 1-function $f_1(x)$ of (1) is computable if and only if the word problem for (1) is solvable and, as was shown above (when $\forall R \in \mathcal{R}$ $|R| > 1$), the Dehn 0-function $f_0(x)$ of (1) is computable if and only if the word problem for (1) is solvable. This shows that $f_0(x)$ is computable if and only if so is $f_1(x)$. $\square$

Proof of Theorem 1.6. Part (a). Let $\mathcal{R}' = (\mathcal{R} \setminus \mathcal{S}) \cup \mathcal{U}$ and $\mathcal{S} = \{S_1, \dots, S_k\}$, $\mathcal{U} = \{U_1, \dots, U_\ell\}$. Since $\mathcal{U} \subset \langle\langle \mathcal{R} \rangle\rangle$, there are van Kampen diagrams Δ_i over the presentation $\langle \mathcal{A} \,\|\, \mathcal{R}' \rangle$ such that $\varphi(\partial\Delta_i) \equiv S_i$, $i = 1, \dots, k$. Denote

$$M_j = \max\{|\Delta_i(j)| \mid i = 1, \dots, k\} \ , \quad j = 0, 1, 2 \ .$$

Consider a van Kampen diagram Δ over (1) and let Π be a face in Δ with $\varphi(\partial\Pi) \equiv S_i^{\varepsilon}$, $\varepsilon = \pm 1$. We replace Π in Δ by a copy of Δ_i if $\varepsilon = 1$ or by a mirror copy of Δ_i if $\varepsilon = -1$. Doing this for all faces Π in Δ with $\varphi(\partial\Pi) \equiv S_i^{\varepsilon}$, where $i = 1, \dots, k$ and $\varepsilon = \pm 1$, results in a diagram Δ' over the presentation $\langle \mathcal{A} \,\|\, \mathcal{R}' \rangle$. Observe that

$$\max(|\Delta(0)|, |\Delta(2)|) \le 2|\Delta(1)| \ , \quad |\Delta'(0)| \le |\Delta(0)| + M_0|\Delta(2)| \ ,$$
$$|\Delta'(1)| \le |\Delta(1)| + M_1|\Delta(2)| \ , \quad |\Delta'(2)| \le |\Delta(2)| + M_2|\Delta(2)| \ .$$

Hence,

$$|\Delta'(0)| \le 2(1+M_0)|\Delta(1)|, \ |\Delta'(1)| \le (1+2M_1)|\Delta(1)|, \ |\Delta'(2)| \le (1+M_2)|\Delta(2)| \ .$$

It follows from these inequalities and the definitions that

$$f_0'(x) \le 2(1+M_0)f_1(x) \ , \ f_1'(x) \le (1+2M_1)f_1(x) \ , \ f_2'(x) \le (1+M_2)f_2(x) \ .$$

Therefore, $f_0'(x) \preccurlyeq f_1(x)$, $f_1'(x) \preccurlyeq f_1(x)$, $f_2'(x) \preccurlyeq f_2(x)$ which, in view of symmetry between $\mathcal{R}$ and $\mathcal{R}'$, imply the desired relations.

Part (b). Let a presentation $\langle \mathcal{A}' \,\|\, \mathcal{R}' \rangle$ be obtained from (1) by a stabilization, $\mathcal{A}' = \mathcal{A} \cup \{b\}$, and $f_j'(x)$, $f_j(x)$, $j = 0, 1, 2$, be their corresponding Dehn functions. Note that if Δ is a diagram over $\langle \mathcal{A}' \,\|\, \mathcal{R}' \rangle$ then every edge e of Δ with $\varphi(e) = b^{\pm 1}$ lies on the boundary $\partial\Delta$ of Δ. This remark and the definitions enable us to conclude that $f_j(x) \le f_j'(x)$ and $f_j'(x) \le f_j(x) + x$, $j = 0, 1, 2$. These inequalities imply the required equivalence $f_j'(x) \simeq f_j(x)$, $j = 0, 1, 2$. $\square$

Proof of Theorem 1.7. Let $\mathcal{R}$ be finite. Then $M = \max\{|R| \mid R \in \mathcal{R}\}$ is also finite and $|\Delta(1)| \le M|\Delta(2)| + |\partial\Delta|/2$ for every diagram Δ over (1), where $|\partial\Delta|$ is the perimeter of Δ. Using the notation of the definition (3), we further have $L_1(W) \le L_2(W) + |W|/2$. Maximizing over all W, where $W \in \langle\langle \mathcal{R} \rangle\rangle$ and $|W| \le x$, we get $f_1(x) \le M f_2(x) + x/2$ and so $f_1(x) \preccurlyeq f_2(x)$. This, together with Theorem 1.4(a), proves the equivalence $f_1(x) \simeq f_2(x)$.

To prove the equivalence $f_0(x) \simeq f_1(x)$, we will argue by induction on the total length $\|\mathcal{R}\| = \sum_{R \in \mathcal{R}} |R|$ of words in $\mathcal{R}$. The base step for $\|\mathcal{R}\| \le 1$ is obvious. If all words in $\mathcal{R}$ have length > 1, then the desired equivalence follows from Theorem 1.4(b). Without loss of generality, we may assume that $\mathcal{R}$ contains a letter $b \in \mathcal{A}$ and the set $\mathcal{R}^b = \mathcal{R} \setminus \{b\}$ is nonempty.

If b, b^{-1} do not occur in words of $\mathcal{R}^b$, then we can apply a stabilization to (1) and obtain the presentation $\langle \mathcal{A} \setminus \{b\} \,\|\, \mathcal{R}^b \rangle$. By the induction hypothesis, its Dehn functions $f_0^b(x)$, $f_1^b(x)$ are equivalent and, by Theorem 1.6(b), $f_j^b(x) \simeq f_j(x)$, $j = 0, 1$. Hence, $f_0(x)$, $f_1(x)$ are also equivalent, as required.

Now assume that b (or b^{-1}) occurs in a relator $R_b \in \mathcal{R}^b$. Let

$$R_b \equiv U b U^{-1} S \ , \tag{15}$$

where $U, S \in F(\mathcal{A})$ and S is cyclically reduced. Clearly, $S \in \langle\langle \mathcal{R}^b \rangle\rangle$. Set

$$\mathcal{R}' = (\mathcal{R} \setminus \{R_b\}) \cup \{S\}$$

and consider the presentation

$$\langle \, \mathcal{A} \,\|\, \mathcal{R}' \, \rangle \ . \tag{16}$$

Note $\|\mathcal{R}'\| < \|\mathcal{R}\|$. Pick a cyclically reduced word $W \in \langle\langle \mathcal{R} \rangle\rangle$ and let Δ be a van Kampen diagram over (1) such that $\varphi(\partial\Delta) \equiv W$ and Δ is minimal relative to $|\Delta(0)|$. If Π is a face in Δ and $\varphi(\partial\Pi) \equiv R_b^{\varepsilon}$, $\varepsilon = \pm 1$, then we consider a subpath p of $\partial\Pi = pq$ whose label $\varphi(p)$ is the subword $U b^{\varepsilon} U^{-1}$ of R_b^{ε} distinguished in (15).

Let $p = uev$, where $\varphi(u) \equiv \varphi(v)^{-1} \equiv U$, $\varphi(e) \equiv b^{\varepsilon}$. If the initial vertex e_- of e were different from its terminal vertex e_+, then we could put two faces π, π', with $\varphi(\partial\pi) = \varphi(\partial\pi')^{-1} = b^{\varepsilon}$, into Δ, see Figure 1, thus making e_-, e_+ merge and decreasing $|\Delta(0)|$ by one. This contradiction to the minimality of $|\Delta(0)|$ shows that $e_- = e_+$.

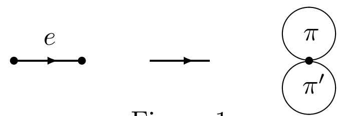

Figure 1

Let $u = u_1u_2$, $v = v_1v_2$ be some factorizations of u, v such that $|u_2| = |v_1|$, where $|u_2|$ denotes the length of u_2, and $(u_2)_- = (v_1)_+$, see Figure 2. We pick such factorizations so that $|u_2|$ is maximal (perhaps, $|u_2| = 0$). Since $(u_2)_- = (v_1)_+$ and $\varphi(u_1) \equiv \varphi(v_2)^{-1}$, we can do the following surgery over Δ. Take the subdiagram bounded by the closed path u_2ev_1 out of Δ and identify the paths u_1 and v_2^{-1}, see Figure 2.

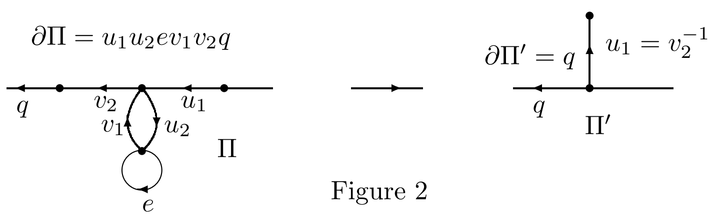

Figure 2

By doing this, we turn the face Π with $\varphi(\partial\Pi) = R_b^{\varepsilon}$ into a face Π' with $\varphi(\partial\Pi') = S^{\varepsilon}$ and do not increase $|\Delta(0)|$. Iterating such surgeries for all faces Π as above, we will obtain a diagram Δ' over the presentation (16) such that $\varphi(\partial\Delta') \equiv W$ and

$$|\Delta'(0)| \leq |\Delta(0)| \ . \tag{17}$$

It follows from the choice of Δ that $L_0(W) = |\Delta(0)|$. Hence, referring to the inequality (17) and the definition (3), we obtain that

$$f_0'(x) \leq f_0(x) \ , \tag{18}$$

where $f_0'(x)$ is the Dehn 0-function of the presentation (16). Since the presentation (16) is obtained from $\langle \mathcal{A} \,\|\, \mathcal{R} \rangle$ by a T-transformation which replaces R_b by S, it follows from Theorem 1.6(a) that $f_1'(x) \simeq f_1(x)$. In view of $\|\mathcal{R}'\| < \|\mathcal{R}\|$, the induction hypothesis applies to (16) and yields that $f_0'(x) \simeq f_1'(x)$. Hence,

$$f_0'(x) \simeq f_1'(x) \simeq f_1(x) \ . \tag{19}$$

By Theorem 1.4(a), $f_0(x) \preccurlyeq f_1(x)$ which, in view of (19), means that $f_0(x) \preccurlyeq f_0'(x)$. This, together with (18), shows that $f_0'(x) \simeq f_0(x)$ and, by (19), we finally have $f_0(x) \simeq f_1(x)$. Theorem 1.7 is proved. $\square$

Let us give examples of group presentations for which the pairs $(f_1(x), f_2(x))$, $(f_0(x), f_1(x))$, $(f_0(x), f_2(x))$ contain nonequivalent functions.

Example 2.1. The presentation

$$\langle\ a, b \ \|\ a^iba^{-i}b^{-1},\ i \in \mathbb{N} \,\rangle$$

defines a free abelian group of rank 2. It is easy to check that $f_1(x) \simeq x^2$ and $f_2(x) \simeq x$. Hence, $f_1(x) \not\simeq f_2(x)$ for this presentation.

Example 2.2. Consider the presentation

$$\langle\ a, b, c \parallel c,\ R_i c^{\ell_i},\ i \in \mathbb{N}\ \rangle\ ,$$

where R_i are words over positive alphabet $\{a, b\}$ that satisfy the small cancelation condition $C'(\lambda)$, $0 < \lambda < 1/6$, see [23], $|R_i| \to \infty$ as $i \to \infty$, and $\ell_i = 2^{|R_i|}$. It is not difficult to verify that $f_0(x) \simeq x$ and $\min(f_1(x), f_2(x)) \geq 2^x$ for every x for which there exists R_i with $|R_i| = x$. Thus, the functions $f_0(x)$, $f_1(x)$ are not equivalent and $f_0(x), f_2(x)$ are not equivalent either.

It is of interest to point out that, when proving Theorem 1.2 (resp. Theorem 1.5), we actually show that the function $L_1(W)$ (resp. $L_0(W)$), where $W \in \langle\langle \mathcal{R} \rangle\rangle$, see (2), is computable if the word problem is solvable for (1). The following example, due to Jockush and Kapovich, gives an indication that Problem 1.1(a) for $j = 2$ might have a negative solution.

Example 2.3. Consider the presentation

$$\langle\ a, b \parallel a^i,\ a^i b^{k_i},\ \ i \in \mathbb{N}\ \rangle\ ,$$

where $\mathbb{K} = \{k_1, k_2, \dots\}$ is a recursively enumerable but not recursive subset of $\mathbb{N}$ with the indicated enumeration and $k_1 = 1$. It is clear that the relator set is decidable and this presentation defines the trivial group, hence the word problem is solvable. On the other hand, it is easy to verify that $L_2(b^k) = 2$, where $k \in \mathbb{N}$, if and only if $k \in \mathbb{K}$. Since $\mathbb{K}$ is not recursive , it follows that the function $L_2(W)$, where W is a word over $\{a^{\pm 1}, b^{\pm 1}\}$, is not computable. It remains to be seen whether this idea would lead to a counterexample to Problem 1.1(a) for $j = 2$.

The following example that gives a negative solution to Problem 1.1(b) for $j = 2$ is due to an anonymous referee.

Example 2.4. Let $\langle \mathcal{A}_0 \parallel \mathcal{R}_0 \rangle$ be a finite presentation with unsolvable word problem and assume that $\mathcal{R}_0$ contains a letter a of $\mathcal{A}_0$. Consider a new letter t, $t \notin \mathcal{A}_0^{\pm 1}$, denote $\mathcal{A} = \mathcal{A}_0 \cup \{t\}$ and let $\mathcal{N}_t$ be the set of all nonempty cyclically reduced words over $\mathcal{A}^{\pm 1}$ that are in the normal closure of t in $F(\mathcal{A})$. Observe that $W \in \mathcal{N}_t$ if and only if W is nonempty, cyclically reduced and $\pi_t(W) = 1$, where $\pi_t : F(\mathcal{A}) \to F(\mathcal{A}_0)$ is the projection homomorphism that erases all occurrences of $t^{\pm 1}$. For every $2k$-tuple $(R_1, X_1, \dots, R_k, X_k)$, where $k \geq 1$, $R_i \in \mathcal{R}_0^{\pm 1}$ and X_i are reduced words over $\mathcal{A}_0^{\pm 1}$ for $i = 1, \dots, k$, we consider the word

$$V(R_1, X_1, \dots, R_k, X_k) = X_1 t R_1 t X_1^{-1} t X_2 t R_2 t X_2^{-1} t \dots X_k t R_k t X_k^{-1} t\ .$$

Let the set $\mathcal{V}_t$ contain the words $V(R_1, X_1, \dots, R_k, X_k)$ for all possible $2k$-tuples $(R_1, X_1, \dots, R_k, X_k)$, as described above. Now we define the presentation

$$\langle\ \mathcal{A} \parallel \mathcal{R} = \mathcal{N}_t \cup \mathcal{V}_t\ \rangle\ . \tag{20}$$

It is clear that $\mathcal{R}$ is a decidable set. Observe that

$$\pi_t(V(R_1, X_1, \dots, R_k, X_k)) = X_1 R_1 X_1^{-1} \dots X_k R_k X_k^{-1}\ , \quad \pi_t(\mathcal{N}_t) = \{1\}\ , \quad t \in \mathcal{N}_t\ .$$

Therefore, a word W over $\mathcal{A}^{\pm 1}$ is in $\langle\langle \mathcal{R} \rangle\rangle$ if and only if $\pi_t(W)$ is in the normal closure $\langle\langle \mathcal{R}_0 \rangle\rangle^{F(\mathcal{A}_0)}$ of the set $\mathcal{R}_0$ in $F(\mathcal{A}_0)$. In particular, the word problem is unsolvable for the presentation (20).

Now assume that $W \in \langle\langle \mathcal{R} \rangle\rangle$. Then $\pi_t(W) \in \langle\langle \mathcal{R}_0 \rangle\rangle^{F(\mathcal{A}_0)}$ and hence there exists a suitable word $V = V(R_1, X_1, \dots, R_k, X_k)$ in $\mathcal{R}$ such that $\pi_t(V) = \pi_t(W)$. Since $\pi_t(WV^{-1}) = 1$, it follows that either $WV^{-1} = 1$ in $F(\mathcal{A})$ or WV^{-1} is conjugate in $F(\mathcal{A})$ to a word in $\mathcal{N}_t$. Writing W in the form $W = (WV^{-1})V$, we see that $L_2(W) \le 2$, where $L_2(W)$ is defined by means of (20). Furthermore, it follows from the definitions that $a \notin \mathcal{R}^{\pm 1}$ which implies that $L_2(a) = 2$. Thus, the Dehn 2-function $f_2(x)$ of the presentation (20) is identically equal to 2 and hence is computable, whereas the word problem for (20) is unsolvable.

We remark that it is also interesting to state Problem 1.1 requiring, in addition, that $\mathcal{R}$ be minimal. The presentations of Examples 2.3–2.4 are far from being minimal and potential counterexamples (if they exist) to the analog of Problem 1.1 with minimal $\mathcal{R}$ could be more difficult to construct.

As was suggested by the referee, it is of interest to consider an "upper bound" form of Problem 1.1 in which the computability of the Dehn function $f_j(x)$ is replaced by the computability of an upper bound of $f_j(x)$. Observe that the upper bound version of Problem 1.1(a) has a straightforward positive solution for both $j = 0, 2$. However, even this relaxed version of Problem 1.1 still has a negative solution for $j = 2$, as follows from Example 2.4.

The referee also pointed out that our idea of the Dehn 1-function f_1 is analogous to the concept of "derivation work" introduced by Birget [3] in 1998 for semigroup and group presentations and, in fact, Birget's "derivation work" function is equivalent to $f_1(x)$ in terms of the linear equivalence $\simeq$. Furthermore, the left-to-right direction of Theorem 1.3 is a consequence of Proposition 3.3 and Corollary 3.4 of [3] where a much stronger constraint for $\mathcal{R}$ is obtained: the membership problem for $\mathcal{R}$ is not only in **NP** but $\mathcal{R}$ can be chosen to be the intersection of two deterministic context-free languages which, in particular, implies that the membership problem for $\mathcal{R}$ is solvable in deterministic linear time.

3. Proofs of Theorem 1.8 and Corollary 1.9

As in Introduction, let Γ denote the group introduced in [7] and defined by presentation (5). This group Γ turned out to possess many interesting properties: Γ is an infinite 3-generator 2-group all of whose nontrivial quotients are finite [7, 11], Γ has bounded width with respect to the lower central series, Γ is of intermediate (between polynomial and exponential) growth [8], Γ is amenable but not elementary amenable [8] etc. A detailed discussion of properties of the group Γ can be found in [8, 11, 12, 17]. This group Γ can be defined in several different ways but, in this article, we will only use the definition of Γ by means of the presentation (5).

Let $\Gamma(0)$ denote the free group $F(a, b, c, d)$ in the alphabet $\{a, b, c, d\}$ and $\mathcal{R}(0)$ be the empty set. Consider the group presentation

$$\Gamma(1) = \langle\ a, b, c, d \parallel a^2,\ b^2,\ c^2,\ d^2,\ bcd,\ (ad)^4\ \rangle\ = \langle\ a, b, c, d \parallel \mathcal{R}(1)\ \rangle\ . \qquad (21)$$

For $i \ge 2$, we define

$$\Gamma(i) = \langle\ a, b, c, d \parallel a^2,\ b^2,\ c^2,\ d^2,\ bcd,\ (ad)^4,\ \sigma^j((ad)^4), \\ \sigma^{j-1}((adacac)^4),\ j = 1, \dots, i-1 \rangle\ =\ \langle\ a, b, c, d \parallel \mathcal{R}(i)\ \rangle\ . \qquad (22)$$

For every $i \in \mathbb{N}^* = \mathbb{N} \cup \{0\} \cup \{\infty\}$, let $\mathcal{N}(i)$ denote the normal closure $\langle\langle \mathcal{R}(i) \rangle\rangle$ of the relator set $\mathcal{R}(i)$ in $F(a, b, c, d)$. It follows from the Lysenok theorem [24]

that $\Gamma = F(a,b,c,d)/\mathcal{N}(\infty)$. Note that, whenever it is not ambiguous, we do not distinguish between a group and its presentation. Unifying the foregoing notation (5), (21), (22), we can define the presentation

$$\Gamma(i) = \langle\ a,b,c,d \parallel \mathcal{R}(i)\ \rangle$$

for every $i \in \mathbb{N}^*$.

Consider the subgroup $\mathcal{H}(i)$ of $\Gamma(i)$ that is generated by the images of the words

$$b,\ c,\ d,\ aba,\ aca,\ ada,\ a^2\ . \tag{23}$$

It is easy to see that $\mathcal{H}(i)$ is of index 2 in $\Gamma(i)$.

Observe that the words (23) are free generators of the free group $\mathcal{H}(0)$. Define a homomorphism

$$\psi_0 : \mathcal{H}(0) \mapsto \Gamma(0) \times \Gamma(0)$$

by setting

$$\psi_0 = \begin{cases} b \mapsto (a,c), \\ c \mapsto (a,d), \\ d \mapsto (1,b), \\ aba \mapsto (c,a), \\ aca \mapsto (d,a), \\ ada \mapsto (b,1) \\ a^2 \mapsto (1,1)\ . \end{cases} \tag{24}$$

Referring to the definition of σ, see (6), we observe that $\sigma(\Gamma(0)) \subseteq \mathcal{H}(0)$. Hence, we can compose

$$\psi_0 \cdot \sigma : \Gamma(0) \mapsto \Gamma(0) \times \Gamma(0)\ .$$

Computing $\psi_0 \cdot \sigma$, we obtain

$$\psi_0 \cdot \sigma = \begin{cases} a \mapsto (d,a), \\ b \mapsto (1,b), \\ c \mapsto (a,c), \\ d \mapsto (a,d)\ . \end{cases} \tag{25}$$

It will be convenient to partition the relator set $\mathcal{R}(\infty)$ of the presentation $\Gamma(\infty)$, or (5), as follows. Let

$$\mathcal{S}(0) = \{a^2,\ b^2,\ c^2,\ d^2,\ bcd,\ (ad)^4,\ (adacac)^4\}$$

and, for $i \geq 1$, we set $\mathcal{S}(i) = \{\sigma^i((ad)^4), \sigma^i((adacac)^4)\}$. Clearly,

$$\mathcal{R}(\infty) = \bigcup_{i=0}^{\infty} \mathcal{S}(i)$$

is a partition of $\mathcal{R}(\infty)$. The *height* $h(R)$ of a relator $R \in \mathcal{R}(\infty)$ is defined to be $j \geq 0$ if $R \in \mathcal{S}(j)$.

Lemma 3.1. *For every $i \in \mathbb{N}^*$, the map ψ_0, defined by (24), extends to a homomorphism*

$$\psi_i : \mathcal{H}(i) \mapsto \Gamma(i) \times \Gamma(i) \tag{26}$$

such that $\ker \psi_\infty = \{1\}$ and $\ker \psi_1 = \mathcal{N}(2)/\mathcal{N}(1)$. In particular, $\ker \psi_1$ is a normal subgroup of $\Gamma(1)$ generated by the images of $\sigma((ad)^4) = (ac)^8$ and $(adacac)^4$.

Proof. Referring to the classic Reidemeister–Schreier rewriting process, see [23] or [27], we conclude that, to prove that ψ_0 extends to a homomorphism (26), it suffices to verify that

$$\psi_0(\mathcal{R}(i)) \cup \psi_0(a\mathcal{R}(i)a) \subset \mathcal{N}(i) \times \mathcal{N}(i) \ . \tag{27}$$

It is easy to check that

$$\psi_0(\mathcal{S}(0) \cup a\mathcal{S}(0)a) \subset \mathcal{N}(1) \times \mathcal{N}(1) \ .$$

In fact, to show this inclusion we do not even need to use the inclusion $(ad)^4 \in \mathcal{N}(1)$. Since $\mathcal{R}(1) \subset \mathcal{S}(0)$, it follows that $\psi_0(\mathcal{R}(1) \cup a\mathcal{R}(1)a) \subset \mathcal{N}(1) \times \mathcal{N}(1)$.

Suppose $U \in \mathcal{R}(i)$, where $i \geq 2$, and $U \notin \mathcal{S}(0)$. It follows from the definitions that $U = \sigma(V^4)$, where V^4 is the σ^j-image of $(ad)^4$ or $(adacac)^4$, where $0 \leq j \leq i-2$, and so $V^4 \in \mathcal{R}(i-2)$. Hence, according to (25),

$$\psi_0(U) = \psi_0(\sigma(V))^4 = (T^4, V^4) \ , \tag{28}$$

where T is a word over the alphabet $\{a, d\}$. Since $a^2, d^2, (ad)^4 \in \mathcal{N}(1)$ and $V^4 \in \mathcal{R}(i-2)$, it follows that $T^4 \in \mathcal{N}(1)$ and $\psi_0(U) \in \mathcal{N}(i) \times \mathcal{N}(i)$, as desired.

It follows from the definition of ψ_0, see (24), that, switching $g \leftrightarrow aga$, where $g \in \{b, c, d\}$, results in switching the first and second components of $\psi_0(e)$. This remark, together with (28) and the equality $\psi_0(a^2) = (1, 1)$, implies that

$$\psi_0(aUa) = \psi_0(a\sigma(V^4)a) = (V^4, T^4) \ . \tag{29}$$

Hence, we also have $\psi_0(aUa) \in \mathcal{N}(i) \times \mathcal{N}(i)$ and the inclusion (27) is proved.

The equality $\ker \psi_\infty = \{1\}$ was observed in the original article [7], see also [17], and the equality $\ker \psi_1 = \mathcal{N}(2)/\mathcal{N}(1)$ is shown in [13]. $\square$

Lemma 3.2. *Let $U \in \mathcal{N}(\infty)$ and $\psi_0(U) = (U_0, U_1)$. Then, in the free group $F(a, b, c, d)$, $U = a\sigma(U_0)a\sigma(U_1)V$, where $V \in \mathcal{N}(2)$.*

Proof. By Lemma 3.1, we have $U_0, U_1 \in \mathcal{N}(\infty)$ and, in view of $\sigma(\mathcal{N}(\infty)) \subset \mathcal{N}(\infty)$, it follows from (25) that

$$\psi_0(\sigma(U_1)) = (T_1, U_1) \ ,$$

where T_1 is a word over $\{a, d\}$, and hence $T_1 \in \mathcal{N}(1)$. As in the proof of Lemma 3.1, in view of the symmetry of the definition of ψ_0, see (29), relative to the switch $g \leftrightarrow aga$, $g \in \{b, c, d\}$, we analogously obtain

$$\psi_0(a\sigma(U_0)a) = (U_0, T_2) \ ,$$

where $T_2 \in \mathcal{N}(1)$. Therefore,

$$\psi_0(a\sigma(U_0)a\sigma(U_1)) = (U_0T_1, T_2U_1) \ .$$

Since $\psi_0(U) = (U_0, U_1)$, it follows from Lemma 3.1 that

$$U^{-1}a\sigma(U_0)a\sigma(U_1) \in \ker \psi_1 \cdot \mathcal{N}(1) = \mathcal{N}(2) \ ,$$

as required. $\square$

Observe that if $R \in \mathcal{S}(i)$, $i \geq 1$, then $\sigma(R) \in \mathcal{S}(i+1)$. On the other hand, if $R \in \mathcal{S}(0)$, then either $\sigma(R) \in \mathcal{S}(1)$ or, up to a cyclic permutation (when $R = bcd$), $\sigma(R) \in \mathcal{S}(0)$ except for the case $R = a^2$. Since $\sigma(a^2) \notin \mathcal{R}(\infty)$, we wish to extend the sets $\mathcal{S}(i)$, $i \geq 1$, by adding $\sigma^i(a^2)$. Set $\mathcal{S}^*(0) = \mathcal{S}(0)$ and define

$$\mathcal{S}^*(i) = \mathcal{S}(i) \cup \{\sigma^i(a^2)\} \quad \text{for} \quad i \geq 1 \ .$$

Note that the map σ, see (6), extends to a homomorphism $\Gamma(\frac{1}{2}) \mapsto \Gamma(\frac{1}{2})$ of the free product $\Gamma(\frac{1}{2}) = \langle\ a, b, c, d \parallel a^2, b^2, c^2, d^2\ \rangle$ of four groups of order 2. Hence,

$\sigma^j(a^2) \in \mathcal{N}(1)$ for every $j \geq 0$. In particular, the addition of the relator $\sigma^j(a^2)$ to $\mathcal{S}(j)$, $j \geq 1$, does not change the normal closure of $\mathcal{S}(j)$. Hence, letting $\mathcal{R}^*(\infty) = \bigcup_{j=0}^{\infty} \mathcal{S}^*(j)$, we obtain another presentation for the group Γ

$$\Gamma^*(\infty) = \langle\ a, b, c, d \parallel \mathcal{R}^*(\infty) = \bigcup_{j=0}^{\infty} \mathcal{S}^*(j)\ \rangle\ , \tag{30}$$

because $\langle\langle \mathcal{R}^*(\infty) \rangle\rangle = \mathcal{N}(\infty)$.

As above, if $R \in \mathcal{S}^*(j)$, then we say that R has the *height* $h(R) = j$.

Assume that W is a word in $\mathcal{N}(\infty)$. As in (4), consider a product for W of the form

$$W = \prod_{j=1}^{L} X_j R_j^{\varepsilon_j} X_j^{-1}\ , \tag{31}$$

where $X_j \in F(a, b, c, d)$, $R_j \in \mathcal{R}^*(\infty)$, and $\varepsilon_j = \pm 1$. An *h^*-tuple* $\tau^W = (\tau_0^W, \tau_1^W, \dots)$ of a word $W \in \mathcal{N}(\infty)$ is defined by means of a product (31) for W so that τ_i^W is the number of factors $X_j R_j^{\varepsilon_j} T_j^{-1}$ in (31) with $h(R_j) = i$. Clearly,

$$f^*_{2,\Gamma}(x) \leq \max_{|W| \leq x} \sum_{j=0}^{\infty} \tau_j^W\ ,$$

where $f^*_{j,\Gamma}(x)$ denotes the Dehn j-function of the presentation $\Gamma^*(\infty)$, see (30), $j = 0, 1, 2$.

Putting the more restrictive inclusion $R_j \in \mathcal{R}(\infty)$ in the foregoing definition, we analogously define an *h-tuple* $\bar\tau^W$ of a word $W \in \mathcal{N}(\infty)$ with respect to the smaller relator set $\mathcal{R}(\infty) \subset \mathcal{R}^*(\infty)$.

Lemma 3.3. *Let $W \in \mathcal{N}(\infty)$ and $|W| \leq x$. Then W possesses an h^*-tuple $\tau^W = (\tau_0^W, \tau_1^W, \dots)$ such that $\tau_i^W = 0$ if $i > \log_2 x$, $\tau_0^W \leq Cx^2$, and $\tau_i^W \leq \frac{C}{2^{i-1}} x^2$ where $1 \leq i \leq \log_2 x$ and $C > 1$ is a constant (independent of W). In particular, the Dehn functions $f^*_{1,\Gamma}(x)$, $f^*_{2,\Gamma}(x)$ of the presentation $\langle\ a, b, c, d \parallel \mathcal{R}^*(\infty)\ \rangle$ of Γ satisfy the following inequalities*

$$f^*_{2,\Gamma}(x) \leq \max_{|W| \leq x} \sum_{j=0}^{\infty} \tau_j^W \leq 3Cx^2\ , \tag{32}$$

$$f^*_{1,\Gamma}(x) \leq \max_{|W| \leq x} \sum_{j=0}^{\infty} 3 \cdot 2^{j+3} \tau_j^W \leq 50Cx^2 \log_2 x\ . \tag{33}$$

Proof. By induction on $x \geq 1$, where W is a word in $\mathcal{N}(\infty)$ with $|W| \leq x$, we will be proving that W has a desired h^*-tuple $\tau^W = (\tau_0^W, \tau_1^W, \dots)$. The base of induction for $x = 1$ is trivial and we assume that $x \geq 2$.

First we note that, using at most $3x$ copies of relators a^2, b^2, c^2, d^2, bcd (when "using" we allow cyclic permutations and inversions), we can turn the word W into a word U over the positive alphabet $\{a, b, c, d\}$ such that $|U| \leq |W| \leq x$ and any cyclic permutation of U contains no subwords of the form a^2 and $g_1 g_2$, where $g_1, g_2 \in \{b, c, d\}$. Indeed, at most x copies of relators a^2, b^2, c^2, d^2 are sufficient to turn W into a positive word W' with $|W'| \leq |W|$. Then, decreasing $|W'|$, we can delete the unwanted subwords by applying $\leq 2x$ additional copies of relators a^2, b^2, c^2, d^2, bcd.

Clearly, $U \in \mathcal{N}(\infty)$. Let $\psi_0(U) = (U_0, U_1)$. By the definitions of ψ_0 and U, we have $\max(|U_0|, |U_1|) \le |U|/2 \le x/2$. It follows from Lemmas 3.1–3.2 that $U_0, U_1 \in \mathcal{N}(\infty)$ and, in the free group $F(a,b,c,d)$,

$$U = a\sigma(U_0)a\sigma(U_1)V \ , \tag{34}$$

where $V \in \mathcal{N}(2)$. Since $|\sigma(U_0)| \le 3|U_0|$, $|\sigma(U_1)| \le 3|U_1|$, we can assume that

$$|V| \le 2 \cdot 3|U_0| + |U| + 2 \le 4x + 2 \ .$$

It is known [13] that the group $\Gamma(2) = F(a,b,c,d)/\mathcal{N}(2)$ contains a subgroup of finite index which is isomorphic to the direct product $F_2 \times F_2$, where F_2 is a free group of rank 2. Moreover, $\Gamma(i) = F(a,b,c,d)/\mathcal{N}(i)$, where $i \ge 2$ is finite, contains a subgroup of finite index isomorphic to the direct product of 2^i copies of F_2, see [13]. Since the equivalence class of Dehn 2-functions (for finite presentations) does not change when passing to subgroups of finite index, and the Dehn 2-function of $F_2 \times F_2$ is quadratic, it follows that there exists a constant $C_0 > 0$ for the presentation $\Gamma(2)$ such that if $Y \in \mathcal{N}(2)$ then Y can be written as a product of at most $C_0|Y|^2$ conjugates of elements of $\mathcal{R}(2)^{\pm 1}$. In particular, the word V can be written as a product of at most

$$C_0|V|^2 \le C_0(4x+2)^2$$

conjugates of elements of $\mathcal{R}(2)^{\pm 1}$. Since $\mathcal{R}(2) \subset \mathcal{S}^*(0) \bigcup \mathcal{S}^*(1)$, it follows that V possesses an h^*-tuple $\tau^V = (\tau_0^V, \tau_1^V, \dots)$ such that

$$\tau_0^V + \tau_1^V \le C_0(4x+2)^2 \quad \text{ and } \quad \tau_i^V = 0 \quad \text{for} \quad i > 1 \ . \tag{35}$$

By the induction hypothesis applied to the words U_0, U_1, we obtain the existence of h^*-tuples $\tau^k = (\tau_0^k, \tau_1^k, \dots)$ for U_k, where $k = 0, 1$, such that $\tau_i^k = 0$ if $i > \log_2(x/2)$ and

$$\tau_0^k \le C(x/2)^2 \ , \qquad \tau_i^k \le \frac{C}{2^{i-1}}(x/2)^2 \tag{36}$$

for i with $1 \le i \le \log_2(x/2)$.

Note that if

$$U_k = \prod_{j=1}^{L_k} X_{jk} R_{jk}^{\varepsilon_{jk}} X_{jk}^{-1} \ ,$$

where $X_{jk} \in F(a,b,c,d)$, $R_{jk} \in \mathcal{R}^*(\infty)$, $\varepsilon_{jk} = \pm 1$, $k = 0, 1$, then

$$\sigma(U_k) = \prod_{j=1}^{L_k} \sigma(X_{jk})\sigma(R_{jk})^{\varepsilon_{jk}}\sigma(X_{jk})^{-1} \ ,$$

where $\sigma(X_{jk}) \in F(a,b,c,d)$. In addition, up to a cyclic permutation (in case $R_{jk} = bcd$), we have that $\sigma(R_{jk}) \in \mathcal{R}^*(\infty)$ with either $h(\sigma(R_{jk})) = h(R_{jk})$, if $R_{jk} \in \{b^2, c^2, d^2, bcd\}$, or $h(\sigma(R_{jk})) = h(R_{jk}) + 1$ otherwise. As was pointed out above, $\sigma(a^2) \notin \mathcal{R}(\infty)$ and this is the reason for extending the set $\mathcal{R}(\infty)$ to $\mathcal{R}^*(\infty)$.

Referring to the presentation (34) for U, we observe that U is a product of words a^2, $a^{-1}\sigma(U_0)a$, $\sigma(U_1)$, V. Hence, an h^*-tuple for U can be obtained as the sum of h^*-tuples for a^2, $a^{-1}\sigma(U_0)a$, $\sigma(U_1)$, V. This observation implies the existence of an h^*-tuple $\tau^U = (\tau_0^U, \tau_1^U, \dots)$ for U with the following properties.

First, since $\tau_{i'}^k = 0$, where $i' > \log_2(x/2)$, $k = 0, 1$, and $\tau_{i'}^V = 0$, where $i' > 1$, it follows that $\tau_i^U \le \tau_{i-1}^0 + \tau_{i-1}^1 + \tau_i^V = 0$ when $i > \log_2 x = \log_2(x/2) + 1 \ge 1$ for

$x \geq 2$. Second, in view of the estimates (35), (36), $x \geq 2$ and the definitions, we have

$$\begin{aligned}
\tau_0^U &\leq 1 + \tau_0^V + \tau_0^0 + \tau_0^1 \leq C_0(4x+2)^2 + 1 + 2C(x/2)^2 \\
&\leq C_0(4x+2)^2 + \frac{C}{2}x^2 + 1 \leq C_0(5x)^2 + 1 + \frac{C}{2}x^2 \ , \\
\tau_1^U &\leq \tau_1^V + \tau_0^0 + \tau_0^1 \leq C_0(4x+2)^2 + 2C(x/2)^2 \leq C_0(5x)^2 + \frac{C}{2}x^2 \ , \quad \text{and} \\
\tau_i^U &\leq \tau_{i-1}^0 + \tau_{i-1}^1 \leq 2 \cdot \frac{C}{2^{i-2}}(x/2)^2 = \frac{C}{2^{i-1}}x^2 \ , \qquad \text{where} \quad i \geq 2 \ .
\end{aligned}$$

Choosing $C \geq 50(C_0+1)$, we obtain that τ_0^U, $\tau_1^U < Cx^2 - 3x$. Recall that U was obtained from W by at most $3x$ applications of copies of relators a^2, b^2, c^2, d^2, bcd. This remark finally proves the existence of an h^*-tuple $\tau^W = (\tau_0^W, \tau_1^W, \dots)$ for W with all of the desired properties.

The inequality (32) is now obvious from the proven estimates for $\tau_0^W, \tau_1^W, \dots$. Note that if $R \in \mathcal{S}^*(i)$ then

$$|R| \leq |\sigma^i((adacac)^4)| = 12 \cdot 2^{i+1} = 3 \cdot 2^{i+3} \ , \tag{37}$$

because every application of σ doubles the length. The inequality (33) also becomes evident. □

Proof of Theorem 1.8: Let $W \in \mathcal{N}(\infty)$ and $|W| \leq x$. According to Lemma 3.3, there exists an h^*-tuple $\tau^W = (\tau_0^W, \tau_1^W, \dots)$ for W with the described properties. Since an h^*-tuple is defined with respect to the relator set $\mathcal{R}^*(\infty)$, we may have relators $\sigma^i(a^2)$ in a presentation for W of the form (31), $1 \leq i \leq \log_2 x$, which are not in $\mathcal{R}(\infty)$. Recall that $\sigma^i(a^2)$ belongs to the normal closure of words a^2, b^2, c^2, d^2. We also note that, as follows from the definition of σ, see (6), $\sigma^i(a^2)$ is a positive word and has length $2^{i+2} - 2$.

Therefore, we can turn $\sigma^i(a^2)$ into the empty word by consecutive deletions of $\frac{2^{i+2}-2}{2} = 2^{i+1} - 1$ subwords of the form g^2, $g \in \{a, b, c, d\}$. This means that $\sigma^i(a^2)$ is a product of $2^{i+1} - 1$ conjugates of relators a^2, b^2, c^2, d^2.

Consequently, replacing each relator $\sigma^i(a^2)$ (or its inverse), $i \geq 1$, by a product of $2^{i+1} - 1$ conjugates of relators a^2, b^2, c^2, d^2 (or their inverses), we can turn a product (31), defined for W by means of the extended set $\mathcal{R}^*(\infty)$, into a similar product, defined for W by means of the original set $\mathcal{R}(\infty)$. Hence, there exists an h-tuple $\tau' = (\tau_0', \tau_1', \dots)$, defined for W by means of $\mathcal{R}(\infty)$, such that

$$\begin{aligned}
\tau_0' &= \tau_0^W + \sum_{i=1}^{[\log_2 x]} (2^{i+1} - 1)\tau_i^W \leq Cx^2 + \sum_{i=1}^{[\log_2 x]} (2^{i+1} - 1) \cdot \frac{C}{2^{i-1}}x^2 \\
&\leq Cx^2 + \sum_{i=1}^{[\log_2 x]} (4 - 2^{1-i})Cx^2 \leq 4Cx^2 \log_2 x \ ,
\end{aligned} \tag{38}$$

$$\tau_i' \leq \tau_i^W \leq \frac{C}{2^{i-1}}x^2 \quad \text{if} \quad 1 \leq i \leq \log_2 x \ , \quad \text{and} \quad \tau_i' = 0 \quad \text{if} \quad i > \log_2 x \ , \tag{39}$$

where $[\log_2 x]$ is the greatest integer ℓ with $\ell \leq \log_2 x$.

Let Δ be a van Kampen diagram over the presentation (5) with $|\partial\Delta| \leq x$. Assuming that Δ is minimal relative to $|\Delta(2)|$, in view of estimates (37), (38), (39)

and Lemma 3.3, we can estimate the numbers $|\Delta(j)|$ of j-cells in Δ, $j=1,2$, and, hence, the Dehn functions $f_{j,\Gamma}(x)$ of the presentation (5) as follows

$$\begin{aligned} f_{1,\Gamma}(x) &\le 2\tau_0' + \sum_{i=0}^{[\log_2 x]} 3\cdot 2^{i+3}\tau_i' + x/2 \le 8Cx^2\log_2 x + 3\cdot 2^4 Cx^2\log_2 x + x/2 \\ &\le 2^6 Cx^2 \log_2 x + x/2\ , \\ f_{2,\Gamma}(x) &\le \sum_{i=0}^{[\log_2 x]} \tau_i' \le 4Cx^2\log_2 x + \sum_{i=1}^{[\log_2 x]} \tau_i \le 4Cx^2\log_2 x + 2Cx^2\ . \end{aligned}$$

Since $f_{1,\Gamma}(1) = f_{2,\Gamma}(1) = 0$ and $C > 1$, it follows from these estimates that

$$f_{1,\Gamma}(x) \le 55Cx^2\log_2 x \quad \text{and} \quad f_{2,\Gamma}(x) \le 6Cx^2\log_2 x\ .$$

By Theorem 1.4(a), $f_{0,\Gamma}(x) \le 2f_{1,\Gamma}(x)$, hence we can pick $C_1 = 110C$ and the proof of Theorem 1.8 is complete. $\square$

Proof of Corollary 1.9: Let W be a word in the normal closure $\langle\langle \mathcal{R}_t\rangle\rangle$ of the relator set $\mathcal{R}_t$ of the presentation (7) of Γ_t. By $|W|_\ell$ denote the number of occurrences of letters ℓ, ℓ^{-1} in W, where $\ell \in \{a,b,c,d,t\}$. Since $tgt^{-1} = \sigma(g)$, where $g \in \{a,b,c,d\}$, it follows from the definitions, see (6), (7), that we can apply relations $tgt^{-1} = \sigma(g)$ to the cyclic word W and eliminate all occurrences of t, t^{-1}. Note that each application of $tat^{-1} = aca$ adds two extra letters and doubles the number of occurrences of a, whereas an application of $tgt^{-1} = \sigma(g)$, where $g \in \{b,c,d\}$, preserves the length. Hence, when eliminating two successive occurrences of t, t^{-1} in the cyclic word W, we get a new cyclic word W_1 such that

$$|W_1|_t = |W|_t - 2\ , \quad |W_1|_a \le 2|W|_a\ , \quad \text{and} \quad |W_1| \le |W| + 2|W|_a\ .$$

Iterating, in $k \ge 1$ steps, we get a word W_k such that

$$\begin{aligned} &|W_k|_t = |W|_t - 2k\ , \quad |W_k|_a \le 2^k|W|_a\ , \quad \text{and} \\ &|W_k| \le |W| + |W|_a(2^1 + \cdots + 2^k) = |W| + |W|_a(2^{k+1}-2)\ . \end{aligned} \tag{40}$$

Making $|W|_t/2$ such steps, we get a word $U = W_{|W|_t/2}$ with no occurrences of t, t^{-1}. Recall that $W \in \langle\langle \mathcal{R}_t\rangle\rangle$, whence $|W|_t$ is even.

Note that, in view of inequalities $|W|_a + |W|_t \le |W|$ and $y < 2^y$, we have

$$|W|_a \cdot 2^{|W|_t/2} \le |W|_a \cdot 2^{(|W|-|W|_a)/2} \le \frac{|W|_a}{2^{|W|_a/2}} \cdot 2^{|W|/2} < 2\cdot 2^{|W|/2}\ . \tag{41}$$

Hence, according to (40), (41),

$$\begin{aligned} |U| &\le |W| + |W|_a \cdot (2^{|W|_t/2+1} - 2) \le |W| + 2|W|_a \cdot 2^{|W|_t/2} \\ &\le 2\cdot 2^{|W|/2} + 4\cdot 2^{|W|/2} = 6\cdot 2^{|W|/2}\ . \end{aligned} \tag{42}$$

It follows from inequalities (40), (41) and $y^2 < 2^3\cdot 2^{y/2}$, where $y > 0$, that the number of relations of the form $tgt^{-1} = \sigma(g)$, where $g \in \{a,b,c,d\}$, that are needed to obtain U from W, does not exceed the following sum

$$\begin{aligned} |W| + |W_1| + \cdots + |W_{|W|_t/2}| &\le \tfrac{1}{2}|W|^2 + |W|_a \cdot \sum_{k=1}^{|W|_t/2} 2^{k+1} \\ &\le 2^2\cdot 2^{|W|/2} + |W|_a \cdot 2^{|W|_t/2+2} \le 12\cdot 2^{|W|/2}\ . \end{aligned} \tag{43}$$

As we saw in the proof of Theorem 1.8, the word U has an h-tuple $\tau' = (\tau'_0, \tau'_1, \dots)$ defined for U by means of $\mathcal{R}(\infty)$ whose entries satisfy the inequalities (38), (39) in which $x = |U|$.

Note that if $R = \sigma^j(V)$, where $V \in \{(ad)^4, (adacac)^4\}$, $j \ge 1$, then there exists a van Kampen diagram E_R over the presentation (7) of Γ_t such that $\varphi(\partial \mathrm{E}_R) \equiv R$, E_R contains a face Π, with $\varphi(\partial\Pi) \equiv V^{-1}$, and Π is surrounded by j annuli each of which consists of faces corresponding to relators $tgt^{-1}\sigma(g)^{-1}$, $g \in \{a, b, c, d\}$, see Figure 3.

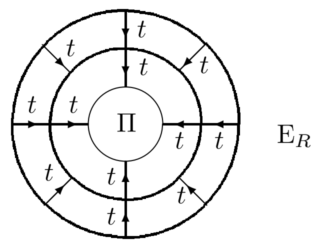

Figure 3

Therefore, in view of (37), the number $|\mathrm{E}_R(2)|$ of 2-cells in E_R can be estimated by

$$|\mathrm{E}_R(2)| \le 1 + \sum_{i=0}^{j-1} 3 \cdot 2^{i+3} < 3 \cdot 2^3 \cdot 2^j \ .$$

Let Δ' be a van Kampen diagram over the presentation (5) such that $\varphi(\partial\Delta') \equiv U$ and the h-tuple $\tau' = (\tau'_0, \tau'_1, \dots)$, defined for U by means of Δ', satisfies the inequalities (38), (39) in which $x = |U|$. As we saw in the proof of Theorem 1.8, any word $U \in \mathcal{N}(\infty)$ has such an h-tuple τ', see (38), (39). Making use of diagrams E_R, we can turn Δ' into a diagram Δ over the presentation (7) of Γ_t so that $\varphi(\partial\Delta) \equiv U$ and

$$\begin{aligned} |\Delta(2)| &\le \tau'_0 + \sum_{j=1}^{[\log_2 x]} 3 \cdot 2^3 \cdot 2^j \cdot \tau'_j \\ &\le 4Cx^2 \log_2 x + 3 \cdot 2^3 \sum_{j=1}^{[\log_2 x]} \frac{2^j}{2^{j-1}} Cx^2 < 2^6 Cx^2 \log_2 x \ . \end{aligned}$$

Hence, there is a diagram Δ_W over (7) such that $\varphi(\partial\Delta_W) \equiv W$ and, in view of inequalities (42), (43),

$$\begin{aligned} |\Delta_W(2)| &\le 12 \cdot 2^{|W|/2} + 2^6 C |U|^2 \log_2 |U| \\ &\le 12 \cdot 2^{|W|/2} + 2^6 \cdot 6^2 C \cdot 2^{|W|} \cdot \log_2(6 \cdot 2^{|W|/2}) \\ &\le C'_2 |W| \cdot 2^{|W|} \ , \end{aligned}$$

where $C'_2 > 1$ is a constant. Since the length of every relator of the presentation (7) does not exceed 24, it follows that

$$|\Delta_W(j)| \le 24|\Delta_W(2)| + \max(|W|, 1) \le \max(25C'_2|W| \cdot 2^{|W|},\ 1)$$

for every $j = 0, 1, 2$. This proves that $f_j(x) \le 25C'_2 x 2^x = C_2 x 2^x$, as required. $\square$

4. Proofs of Theorem 1.10 and Corollary 1.11

Proof of Theorem 1.10: Recall that $n \geq 2^{48}$ is a fixed integer, n is either odd or divisible by 2^9 and $m \geq 2$. Under these assumptions, lemmas of article [19] apply to diagrams over the presentation $B(m,n,\infty) = \langle \mathcal{A} \parallel \mathcal{R}_B \rangle$, see (9), and yield that, for every $i \geq 1$, the word A_i exists and the limit group, defined by $B(m,n,\infty)$, is naturally isomorphic to the free m-generator Burnside group $B(m,n) = F(\mathcal{A})/F(\mathcal{A})^n$, where $F(\mathcal{A})$ is the free group over $\mathcal{A}$.

Let W be a word in the normal closure $\langle\langle \mathcal{R}_B \rangle\rangle$ of $\mathcal{R}_B$. Then there exists a reduced diagram Δ over $B(m,n,\infty)$, and hence over $B(m,n,i_0)$ for some $i_0 \geq 0$, see (8), such that $\varphi(\partial\Delta) \equiv W$. Recall that the boundary $\partial\Delta$ of a van Kampen (or disk) diagram Δ is oriented clockwise and the boundary $\partial\Pi$ of a face of Δ is oriented counterclockwise. Here and below we are using the notation and terminology of article [19] and the reader is referred to [19] for more details.

By induction on the perimeter $|\partial\Delta|$ of a reduced diagram Δ over $B(m,n,i_0)$, we will be proving that the number $|\Delta(1)|$ of 1-cells in Δ does not exceed $|\partial\Delta|^{19/12}$. Since this claim is trivial when Δ contains no 2-cells, in which case $|\Delta(1)| = |\partial\Delta|/2$, we may assume that Δ contains a 2-cell.

Denote $w = \partial\Delta$, $x = |\partial\Delta|$ and consider a factorization $w = w_1 w_2 \dots w_8$, where for every $i = 1, \dots, 8$

$$x/8 - 1 = |w|/8 - 1 < |w_i| < |w|/8 + 1 = x/8 + 1 \ , \tag{44}$$

where $|p|$ denotes the length of a path p.

By Lemmas 5.7, 9.8 [19], we can find a face Π and a system of subdiagrams $\Gamma_1, \dots, \Gamma_8$ in Δ, see Figure 4, such that Γ_i is a contiguity subdiagram between Π and w_i and

$$\sum_{i=1}^{8} |\Gamma_i \wedge \partial\Pi| > \theta |\partial\Pi| \ , \quad \text{where} \quad \theta = 0.99 \ . \tag{45}$$

We remark that, for some i, Γ_i might not be defined. However, in view of (45), at least one of $\Gamma_1, \dots, \Gamma_8$ is defined.

Let $\partial\Gamma_i = b_i u_i c_i q_i$ be the standard boundary of Γ_i (if Γ_i is defined), where $u_i = \Gamma_i \wedge w_i$, $q_i = \Gamma_i \wedge \partial\Pi$, and let $w_i = r_i u_i s_i$, $i = 1, \dots, 8$, see Figure 4. If both Γ_i and Γ_{i+1} are defined, here and below indices are considered mod 8, then we let $q_{i+1} t_i q_i$ be a subpath of $\partial\Pi$ and Δ_i denote the subdiagram of Δ bounded by the path $\partial\Delta_i = c_i^{-1} s_i r_{i+1} b_{i+1}^{-1} t_i$, see Figure 4. Informally, Δ_i sits between Γ_i and Γ_{i+1} in the annulus $\Delta - \operatorname{Int}\Pi$ and t_i sits between q_{i+1} and q_i in the cycle $\partial\Pi$.

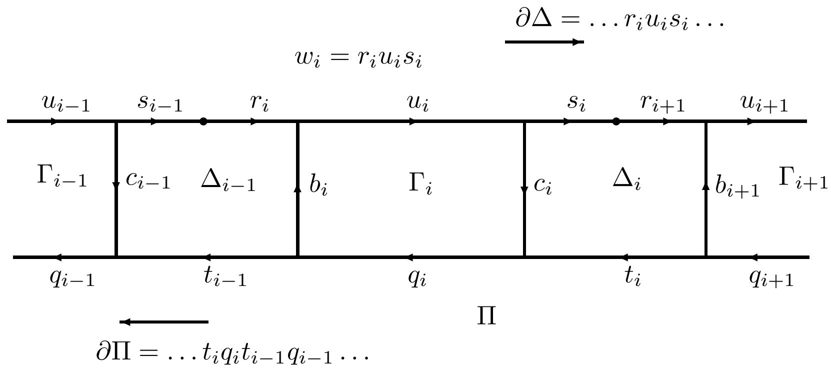

Figure 4

It follows from Lemmas 3.1, 9.8 [19] applied to Γ_i that

$$\max(|b_i|, |c_i|) < \gamma|\partial\Pi| \ , \quad \text{where} \quad \gamma = 2^{-33} \ , \quad i = 1, \dots, 8 \ . \tag{46}$$

By Lemmas 6.1, 9.8 [19] applied to Γ_i, we also have that

$$\rho|q_i| \leq |u_i| + |b_i| + |c_i| \ , \quad \text{where} \quad \rho = 0.95 \ , \quad i = 1, \dots, 8 \ . \tag{47}$$

It follows from Lemmas 6.2, 9.8 [19] applied to the diagram Δ and its face Π that

$$\rho n \leq \rho|\partial\Pi| \leq |\partial\Delta| = x \ . \tag{48}$$

Similar to Storozhev's arguments [30, Section 28.2], we consider two cases which are whether or not all of the subdiagrams $\Gamma_1, \dots, \Gamma_8$ are actually defined.

First we assume that at least one of $\Gamma_1, \dots, \Gamma_8$ is missing. Reindexing if necessary, suppose that Γ_1, Γ_j are present and $\Gamma_2, \dots, \Gamma_{j-1}$ are missing, where $j-1 \in \{3, \dots, 8\}$. Let us emphasize that it is possible that $j-1 = 8$ and $j = 1$, that is, the only contiguity subdiagram, which is defined, is Γ_1.

If t is a path in Δ, then t_-, t_+ denote the initial, terminal, respectively, vertices of t. Consider a path $p = c_1 v^{-1} b_j$ in Δ, where v is the subpath of $\partial\Pi^{-1}$ that connects the vertices $(c_1)_+$, $(b_j)_-$ and contains no paths q_i, $i = 1, \dots, 8$. By estimates (46),

$$|p| < (1 - \theta + 2\gamma)|\partial\Pi| \ . \tag{49}$$

The vertices $p_-, p_+ \in \partial\Delta$ define a factorization $\partial\Delta = y_1 y_2$, where y_2 contains the subpath u_1 of $\partial\Delta$. Cutting Δ along p, we obtain two diagrams E_1, E_2 with $\partial\mathrm{E}_1 = y_1 p^{-1}$, $\partial\mathrm{E}_2 = p y_2$. Since y_1 contains w_2, it follows from (44) and (48) that

$$|y_1| \geq x/8 - 1 \geq (1/8 - (\rho n)^{-1})x \geq (\rho/8 - n^{-1})|\partial\Pi| \ . \tag{50}$$

Adding up all estimates (47) (for those Γ_i that are defined), in view of (45) and (46), we have

$$\rho\theta|\partial\Pi| < \sum_{i=1}^{8} |u_i| + 14\gamma|\partial\Pi| \leq |y_2| + 14\gamma|\partial\Pi| \ ,$$

because every $u_i = \Gamma_i \wedge w_i = \Gamma_i \wedge \partial\Delta$ is a subpath of y_2. Therefore,

$$|y_2| > (\rho\theta - 14\gamma)|\partial\Pi| \ . \tag{51}$$

Now we see from (49), (50), (51) that

$$|p| < (1-\theta+2\gamma)\min\{(\rho/8-n^{-1})^{-1}|y_1|,\ (\rho\theta-14\gamma)^{-1}|y_2|\} \\ < 0.1\min(|y_1|,|y_2|)\ . \tag{52}$$

In particular, $|\partial \mathrm{E}_1|, |\partial \mathrm{E}_2| < x$. Since $\mathrm{E}_1, \mathrm{E}_2$ are reduced diagrams, the induction hypothesis applies to $\mathrm{E}_1, \mathrm{E}_2$ and yields that $|\mathrm{E}_k(1)| \le |\partial \mathrm{E}_k|^{19/12}$, $k = 1, 2$. Since $|\Delta(1)| \le |\mathrm{E}_1(1)| + |\mathrm{E}_2(1)|$, it follows that

$$\begin{aligned} |\Delta(1)| &\le (|p|+|y_1|)^{19/12} + (|p|+|y_2|)^{19/12} \\ &\le \max(|p|+|y_1|, |p|+|y_2|)^{7/12}(|y_1|+|y_2|+2|p|) \\ &\le \max(|p|+|y_1|, |p|+|y_2|)^{7/12}(x+2|p|)\ . \end{aligned} \tag{53}$$

In view of estimate (52),

$$\begin{aligned} \max(|p|+|y_1|, |p|+|y_2|) &= \max(x-|y_2|+|p|, x-|y_1|+|p|) \\ &\le x - 10|p| + |p| = x - 9|p|\ . \end{aligned} \tag{54}$$

By inequalities (48) and (49),

$$|p| < \rho^{-1}(1-\theta+2\gamma)x < 0.02x\ . \tag{55}$$

Let $r = \frac{|p|}{x}$. Then $0 < r < 0.02$ and we can derive from (53), (54), (55) that

$$\begin{aligned} |\Delta(1)| &< (1-9r)^{7/12}(1+2r)x^{19/12} < (1-9r)^{1/2}(1+2r)x^{19/12} \\ &\le ((1-9r)(1+2r)^2)^{1/2}x^{19/12} < ((1-9r)(1+6r))^{1/2}x^{19/12} < x^{19/12}\ , \end{aligned}$$

as required.

Now suppose that all of Γ_i, $i = 1, \dots, 8$, are defined. Taking $\operatorname{Int}\Pi$ out of Δ, we get an annular diagram Δ_0. Cutting Δ_0 along all paths b_i, c_i, we obtain 16 disk diagrams Γ_i, Δ_i, $i = 1, \dots, 8$, see Figure 4. It follows from the definitions that

$$|\partial\Gamma_i| = |b_i| + |u_i| + |c_i| + |q_i|\ , \qquad |\partial\Delta_i| \le |w_i| + |w_{i+1}| + |c_i| + |b_{i+1}| + |t_i|\ .$$

Hence, in view of estimates (44), (45), (46), (47), (48), we further have

$$\begin{aligned} |\partial\Gamma_i| &< (1+\rho^{-1})(|u_i| + 2\gamma|\partial\Pi|) \\ &\le (1+\rho^{-1})(1/8 + (\rho n)^{-1} + 2\gamma\rho^{-1})x < 0.258x\ , \end{aligned} \tag{56}$$

$$\begin{aligned} |\partial\Delta_i| &\le 2(1/8 + (\rho n)^{-1})x + (1-\theta+2\gamma)|\partial\Pi| \\ &\le (1/4 + 2(\rho n)^{-1} + \rho^{-1}(1-\theta+2\gamma))x < 0.262x\ . \end{aligned} \tag{57}$$

Since Γ_i, Δ_i, $i = 1, \dots, 8$, are reduced diagrams, the induction hypothesis applies to them and yields that $|\Gamma_i(1)| \le |\partial\Gamma_i|^{19/12}$, $|\Delta_i(1)| \le |\partial\Delta_i|^{19/12}$, $i = 1, \dots, 8$. Therefore,

$$\begin{aligned} |\Delta(1)| &\le \sum_{i=1}^{8}(|\Gamma_i(1)| + |\Delta_i(1)|) \le \sum_{i=1}^{8}(|\partial\Gamma_i|^{19/12} + |\partial\Delta_i|^{19/12}) \\ &\le \max(|\partial\Gamma_i|, |\partial\Delta_i|)^{7/12} \cdot \sum_{i=1}^{8}(|\partial\Gamma_i| + |\partial\Delta_i|)\ . \end{aligned} \tag{58}$$

It follows from the definitions and estimates (46), (48) that

$$\sum_{i=1}^{8}(|\partial\Gamma_i| + |\partial\Delta_i|) = |w| + |\partial\Pi| + \sum_{i=1}^{8}(|b_i| + |c_i|) \leq (1+\rho^{-1})x + 32\gamma|\partial\Pi|$$
$$\leq (1+\rho^{-1}+32\gamma\rho^{-1})x < 2.06x \ . \tag{59}$$

In view of estimates (56), (57), (59), it follows from (58) that

$$|\Delta(1)| \leq (0.262x)^{7/12}\cdot 2.06x < x^{19/12} \ ,$$

as desired. Thus the desired upper bounds $f_{1,B}(x) \leq x^{19/12}$, $f_{1,B(i_0)}(x) \leq x^{19/12}$ are proven. The inequalities $f_{0,B}(x) \leq 2x^{19/12}$, $f_{0,B(i_0)}(x) \leq 2x^{19/12}$ now follow from Theorem 1.4(a) and the inequalities $f_{2,B}(x) \leq \frac{2}{n}x^{19/12}$, $f_{2,B(i_0)}(x) \leq \frac{2}{n}x^{19/12}$ hold because $|R| \geq n$ for every relator R of $B(m,n,\infty)$, $B(m,n,i_0)$. The proof of Theorem 1.10 is complete. □

Proof of Corollary 1.11: Let W be a word over the alphabet $\mathcal{A}^{\pm 1}$. Then W represents the identity element of the group $B(m,n) = F(\mathcal{A})/F(\mathcal{A})^n$, defined by $B(m,n,\infty)$, if and only if there exists a van Kampen diagram Δ_W over $B(m,n,\infty)$ with $\varphi(\partial\Delta_W) \equiv W$. By Theorem 1.10, we can assume that $|\Delta_W(1)| \leq |W|^{19/12}$. As in the proof of Theorem 1.3, this polynomial bound enables us to take a van Kampen diagram Δ with $|\Delta(1)| \leq |W|^{19/12}$ as a certificate to verify whether $W = 1$ in $B(m,n)$. In polynomial time with respect to $|W|$, we can certify that $\varphi(\partial\Delta) \equiv W$ and, for every face Π in Δ, the label $\varphi(\partial\Pi)$ is the nth power of a word. If these conditions are satisfied, then $W = 1$ in $B(m,n)$, as required.

Now let U, V be two words over the alphabet $\mathcal{A}^{\pm 1}$. It follows from Lemmas 6.3, 9.2 [19] and Theorem 1.10 that U, V are conjugate in $B(m,n)$ if and only if there exists an annular diagram $\Delta_{U,V}$ over $B(m,n,\infty)$ such that the oriented components p, q of the boundary $\partial\Delta_{U,V}$ are labelled by the cyclic words U, V^{-1} and

$$|\Delta_{U,V}(1)| \leq (1.01(|U|+|V|))^{19/12} \leq 1.02(|U|+|V|)^{19/12} \ .$$

Hence, as above, an annular diagram Δ with $|\Delta(1)| \leq 1.02(|U|+|V|)^{19/12}$ can be used to certify, in polynomial time of $|U|+|V|$, that U and V are conjugate in $B(m,n)$ by checking that $\varphi(p) \equiv U$, $\varphi(q) \equiv V^{-1}$ and that, for every face Π in Δ, $\varphi(\partial\Pi)$ is the nth power of a word. □

Acknowledgements: We wish to thank M. Sapir for helpful discussions and for pointing out to the notion of the verbal Dehn function and to other results of [32]. We also thank C. Jockush and I. Kapovich for providing Example 2.3. We are indebted to P. de la Harpe for many useful suggestions and to the anonymous referee for a number of valuable remarks and for providing Example 2.4.

References

[1] S. I. Adian, *The Burnside problem and identities in groups*, Nauka, Moscow, 1975; English translation: Springer-Verlag, 1979.

[2] G. Baumslag, C. F. Miller III, and H. Short, *Isoperimetric inequalities and the homology of groups*, Invent. Math. **113**(1993), 531–560.

[3] J.-C. Birget, *Infinite string rewrite systems and complexity*, J. Symbolic Comp. **25**(1998) 759–793.

[4] S. G. Brick, *On Dehn functions and products of groups*, Trans. Amer. Math. Soc. **335**(1993), 369–384.

[5] J.-C. Birget, A. Yu. Ol'shanskii, E. Rips, and M. V. Sapir, *Isoperimetric functions of groups and computational complexity of the word problem*, Ann. of Math. **156**(2002), 467–518.
[6] M. R. Bridson and N. Brady, *There is only one gap in the isoperimetric spectrum*, Geom. Funct. Anal. **10**(2000) 1053–1070.
[7] R. I. Grigorchuk, *On the Burnside problem for periodic groups*, Funct. Anal. Appl. **14**(1980), 53–54.
[8] R. I. Grigorchuk, *The growth degrees of finitely generated groups and the theory of invariant means*, Izv. Akad. Nauk. SSSR. Ser. Mat. **48**(1984), 939–985.
[9] R. I. Grigorchuk, *An example of a finitely presented amenable group that does not belong to the class EG*, Mat. Sbornik **189**(1998), 79–100.
[10] R. I. Grigorchuk, *On the system of defining relations and the Schur multiplier of periodic groups generated by finite automata*, in *Groups St. Andrews 1997 in Bath, I*, London Math. Soc. Lecture Note Ser. **260**, 1999, 290–317.
[11] R. I. Grigorchuk, *Just infinite branch groups*, in *New horizons in pro-p-groups*, Progr. Math. **184**, Birkhäuser, 2000, 121–179.
[12] R. I. Grigorchuk, *Solved and unsolved problems around one group*, in *Infinite groups: geometric, combinatorial and dynamical aspects*, Progr. Math. **248**, Birkhäuser, 2005, 117–218.
[13] R. I. Grigorchuk and P. de la Harpe, *Limit behaviour of exponential growth rates for finitely generated groups*, In *Essays on geometry and related topics*, Monogr. Enseign. Math. **38**, Enseignement Math., 2001, 351–370.
[14] M. Gromov, *Hyperbolic groups*, In *Essays in Group Theory* (S. Gersten, ed.), MSRI Publ. 8, Springer-Verlag, 1987, 75–263.
[15] V. S. Guba, *The Dehn function of Richard Thompson's group F is quadratic*, Invent. Math. **163**(2006), 313–342.
[16] V. S. Guba and M. V. Sapir, *On Dehn functions of free products of groups*, Proc. Amer. Math. Soc. **127**(1999), 1885–1891.
[17] P. de la Harpe, *Topics in geometric group theory*, Univ. Chicago Press, 2000.
[18] S. V. Ivanov, *On the Burnside problem on periodic groups*, Bull. Amer. Math. Soc. **27**(1992), 257–260.
[19] S. V. Ivanov, *The free Burnside groups of sufficiently large exponents*, Internat. J. Algebra Comp. **4** (1994), 1–308.
[20] S. V. Ivanov, *On the Burnside problem for groups of even exponent*, Documenta Mathematica, Extra Volume ICM-98 **II**(1998), 67–76.
[21] S. V. Ivanov, Embedding free Burnside groups in finitely presented groups, *Geom. Dedicata* **111**(2005), 87–105.
[22] S. V. Ivanov, *On balanced presentations of the trivial group*, Invent. Math. **165**(2006), 525–549.
[23] R. C. Lyndon and P. E. Schupp, *Combinatorial group theory*, Springer-Verlag, 1977.
[24] I. G. Lysenok, *A set of defining relations for the Grigorchuk group*, Matem. Zametki **38**(1985), 503–516.
[25] I. G. Lysenok, *Infinite Burnside groups of even period*, Izv. Ross. Akad. Nauk Ser. Mat. **60**(1996), 3–224.
[26] K. Madlener, F. Otto, *Pseudo-natural algorithms for the word problem for finitely presented groups*, J. Symbolic Comp. **1**(1985), 383–418.
[27] W. Magnus, J. Karras, D. Solitar, *Combinatorial group theory*, Interscience Publ., 1966.
[28] P. S. Novikov and S. I. Adian, *On infinite periodic groups I, II, III*, Izv. Akad. Nauk SSSR Ser. Mat. **32**(1968), 212–244, 251–524, 709–731.
[29] A. Yu. Ol'shanskii, *On the Novikov-Adian theorem*, Mat. Sbornik **118**(1982), 203–235.
[30] A. Yu. Ol'shanskii, *Geometry of defining relations in groups*, Nauka, Moscow, 1989; English translation: *Math. and Its Applications, Soviet series*, vol. 70, Kluwer Acad. Publ., 1991.
[31] A. Yu. Ol'shanskii, *Hyperbolicity of groups with subquadratic isoperimetric inequality*, Internat. J. Algebra Comput. **1**(1991), 281–289.
[32] A. Yu. Ol'shanskii and M. V. Sapir, *Embeddings of relatively free groups into finitely presented groups*, Contemp. Math. **264**(2000), 23–47.
[33] A. Yu. Ol'shanskii and M. V. Sapir, *Length and area functions on groups and quasi-isometric Higman embeddings*, Internat. J. Algebra Comput. **11**(2001), 137–170.
[34] A. Yu. Ol'shanskii and M. V. Sapir, *Non-amenable finitely presented torsion-by-cyclic groups*, Publ. Math. IHES **96**(2002), 43–169.

[35] A. Yu. Ol'shanskii and M. V. Sapir, *Groups with small Dehn functions and bipartite chord diagrams*, Geom. Funct. Anal. **16**(2006), 1324–1376.

[36] C. H. Papadimitriou, *Computational complexity*, Addison-Wesley Publ., 1994.

Department of Mathematics, Mailstop 3368, Texas A&M University, College Station, TX 77843-3368, USA

E-mail address: `grigorch@math.tamu.edu`

Department of Mathematics, University of Illinois, 1409 West Green Street, Urbana, IL 61801, USA

E-mail address: `ivanov@math.uiuc.edu`